



CC cell or organism of interest. Unlike other markers used in protein  
 CC labeling, such as beta-galactosidase and Luciferase, fluorescent proteins  
 CC do not require an exogenous cofactor or substrate. Methods involving  
 CC fluorescent proteins are also less laborious and less difficult to  
 CC control than the traditional methods of fluorescent labeling, where a  
 CC protein of interest is purified and then covalently conjugated to a  
 CC fluorophore derivative. Novel fluorescent proteins isolated from species  
 CC of the class Anthozoa can be used as markers for gene expression and  
 CC protein localization studies, and in fluorescence resonance energy  
 CC transfer (FRET) reactions. They may have improved properties and better  
 CC suitability for larger excitations compared to prior art fluorescent  
 CC proteins such as green fluorescent protein. (Updated on 12-SEP-2003 to  
 CC standardise OS field)

CC Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 3; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MRSSKNVKEEMRFKVRMEGTVNGHFEIEBEGSGRPYEGHNTVYKLYTKGGLPFPAMD 60  
 1 MRSSKNVKEEMRFKVRMEGTVNGHFEIEBEGSGRPYEGHNTVYKLYTKGGLPFPAMD 60  
 QY 61 LSPQFQYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVYTQDSSLQDGCFTY 120  
 DB 61 LSPQFQYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVYTQDSSLQDGCFTY 120  
 QY 121 KYKFTGVNFPDGPVMQKTMGWEASTERTLYPRDGVYKGEIHKALKDKDGGHYLVEFKSI 180  
 DB 121 KYKFTGVNFPDGPVMQKTMGWEASTERTLYPRDGVYKGEIHKALKDKDGGHYLVEFKSI 180  
 QY 181 YMAKRPVQLPGYIYVDSKLDITSHNEDYITVEOYERTBGRHHLFL 225  
 DB 181 YMAKRPVQLPGYIYVDSKLDITSHNEDYITVEOYERTBGRHHLFL 225

#### RESULT 2

AAB01622 standard; protein; 225 AA.

AC AAB01622;  
 XX  
 DT 12-DEC-2000 (first entry)  
 XX  
 DE Discosoma sp. red fluorescent protein drFP583.  
 XX  
 KW Anthozoa; fluorescent protein; fluorescence labeling; amFP486; cFP484;  
 KM zFP506; drFP583; drFP483; asFP600; dgFP512; dmFP592.  
 XX  
 OS Discosoma sp.  
 XX  
 FN WO200034526-A1.  
 XX  
 PD 15-JUN-2000.  
 XX  
 PF 10-DEC-1999; 99WO-US029405.  
 XX  
 PR 11-DEC-1998; 98US-00210330.  
 XX  
 PA (CLON-) CLONTECH LAB INC.  
 XX  
 PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV;  
 XX  
 DR WPI; 2000-423451/36.  
 XX  
 PT Novel method for identifying a DNA sequence encoding fluorescent proteins  
 PT from non-bioluminescent Anthozoa which are useful for fluorescent  
 PT labeling and as markers.  
 XX  
 PS Claim 3; Page 68-69; 73pp; English.  
 XX  
 CC The present sequence is Discosoma sp. red fluorescent protein drFP583. It

CC was isolated using the Aequorea victoria green fluorescent protein (GFP)  
 CC sequence, which was used to design PCR primers which might isolate other  
 CC fluorescent proteins from a number of species of Anthozoa. These were  
 CC Anemonia majano, Clavularia sp., Zoanthus sp., Discosoma sp. and Anemonia  
 CC sulcata. The cDNA obtained was then screened in the search for sequences  
 CC encoding fluorescent proteins. The other proteins found in this manner  
 CC were cFP484, zFP506, drFP583, amFP486, drFP600, dgFP512 and  
 CC dmFP592. These proteins can be used as fluorescent labels (for gene  
 CC expression and protein localisation studies and in fluorescence resonance  
 CC energy transfer (FRET) studies) in place of fluorophore derivatives and  
 CC luciferases, as these involve laborious processes and the latter require  
 CC cofactors. They can also be used in place of GFP, which is too stable to  
 CC be useful when studying short-term or repetitive events

CC Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 3; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEEMRFKVRMEGTVNGHFEIEBEGSGRPYEGHNTVYKLYTKGGLPFPAMD 60  
 DB 1 MRSSKNVKEEMRFKVRMEGTVNGHFEIEBEGSGRPYEGHNTVYKLYTKGGLPFPAMD 60  
 QY 61 LSPQFQYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVYTQDSSLQDGCFTY 120  
 DB 61 LSPQFQYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVYTQDSSLQDGCFTY 120  
 QY 121 KYKFTGVNFPDGPVMQKTMGWEASTERTLYPRDGVYKGEIHKALKDKDGGHYLVEFKSI 180  
 DB 121 KYKFTGVNFPDGPVMQKTMGWEASTERTLYPRDGVYKGEIHKALKDKDGGHYLVEFKSI 180  
 QY 181 YMAKRPVQLPGYIYVDSKLDITSHNEDYITVEOYERTBGRHHLFL 225  
 DB 181 YMAKRPVQLPGYIYVDSKLDITSHNEDYITVEOYERTBGRHHLFL 225

#### RESULT 3

AAG65509 standard; protein; 225 AA.

AC AAG65509;  
 XX  
 DT 30-NOV-2001 (first entry)  
 XX  
 DE Anthozoan red fluorescent protein sequence.  
 XX  
 KW Fluorescent protein; Anthozoan; fluorescence; marker; FRET; red.  
 XX  
 OS Anthozoa.  
 XX  
 FN WO200162919-A1.  
 XX  
 PD 30-AUG-2001.  
 XX  
 PF 13-FEB-2001; 2001WO-US004625.  
 XX  
 PR 23-FEB-2000; 2000US-0184732P.  
 XX  
 PA (AURO-) AURORA BIOSCIENCES CORP.  
 XX  
 PI Nelson D, Zamatra E, Tsien R;  
 XX  
 DR WPI; 2001-557704/62.  
 XX  
 PT Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise  
 PT functional red fluorescent proteins, and the encoding nucleic acids, with  
 PT key mutations for improving the proteins function.  
 XX  
 PS Disclosure; Page 85; 90pp; English.  
 XX  
 CC The invention provides a nucleic acid encoding functional red fluorescent  
 CC protein (II) that differs from the sequence of an Anthozoan red



CC fluorescent protein by at least one amino acid substitution, and with  
CC different fluorescent properties. The red fluorescent protein of the  
CC invention can be expressed by standard recombinant methodology. (II) are  
CC used a fluorescent markers and FRRT partners. It is used for identifying  
CC protein-protein interactions. (II) is also suitable for multiplexed  
CC fluorescent analysis and FRRT-based applications using existing Aequorea  
CC fluorescent proteins. (II) has improved brightness, reduced spectral  
CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
CC The key mutations in the encoding nucleic acids provide improved folding,  
CC brightness, and create (II) with sharper, more defined excitation and  
CC emission peaks when expressed in mammalian cells. The present sequence  
CC represents an anthozoan fluorescent protein

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSKVIVEFMRFKYRMETVNGHFEIIEGEGRPYEGHNTVKLKVTGGPLPAMD 60  
DB 1 MSSSKVIVEFMRFKYRMETVNGHFEIIEGEGRPYEGHNTVKLKVTGGPLPAMD 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFBGCVTVTODSSIQDCGFI 120  
DB 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFBGCVTVTODSSIQDCGFI 120  
QY 121 KYKFTGVNPPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGHYLVEPKSI 180  
DB 121 KYKFTGVNPPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGHYLVEPKSI 180  
QY 181 YMAKKFVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKFVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 4

ID ABB08834 standard; protein; 225 AA.

AC ABB08834;

DT 29-MAY-2002 (first entry)

DE Yeast optimised RFP SEQ ID NO 17.

XX Yeast: red fluorescent protein; RFP; plant; transgenic; GFP;  
XX Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
XX Escherichia coli; green fluorescent protein; biotechnology.

OS Anthozoa.

PN DE20001395-U1.

PD 15-MAR-2001.

PF 27-JAN-2000; 2000DE-02001395.

PR 27-JAN-2000; 2000DE-02001395.

PA (GFCB-) GFC BIOTECH AG.

DR WPI: 2002-228394/29.

DR N-PSDB; ABA95905, ABA95921, ABA95922.

XX New DNA encoding red fluorescent protein, useful as marker in  
XX biotechnology, has sequence optimized for expression in eukaryotes,  
XX especially yeast or plants.

PS Disclosure, Page 13-14; 19pp; German.

XX The invention relates to DNA (I) containing either sequence ABA95905 or  
XX sequence ABA95906 encoding a yeast optimised Red fluorescent Protein

CC (YFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of the yeast optimised RFP

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSKVIVEFMRFKYRMETVNGHFEIIEGEGRPYEGHNTVKLKVTGGPLPAMD 60  
DB 1 MSSSKVIVEFMRFKYRMETVNGHFEIIEGEGRPYEGHNTVKLKVTGGPLPAMD 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFBGCVTVTODSSIQDCGFI 120  
DB 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFBGCVTVTODSSIQDCGFI 120  
QY 121 KYKFTGVNPPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGHYLVEPKSI 180  
DB 121 KYKFTGVNPPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGHYLVEPKSI 180  
QY 181 YMAKKFVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKFVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 5

ID AAE28833 standard; protein; 225 AA.

AC AAE28833;

DT 27-DEC-2002 (first entry)

DE Discosoma sp. drFP583 (NFP-6) wild-type protein.

XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
XX fluorescence activated cell sorting application; fluorescent timer;  
XX biosensor; fluorescence resonance energy transfer application; FRRT;  
XX colouring agent; recombinant DNA application; analyte detection assay;  
XX sunscreen; second messenger detector; drFP583 protein; NFP-6.

OS Discosoma sp.

PN WO200268459-A2.

PD 06-SEP-2002.

PF 20-FEB-2002; 2002WO-US005749.

PR 21-FEB-2001; 2001US-0270983P.

PR 04-DEC-2001; 2001US-00006922.

PA (CLON-) CLONTECH LAB INC.

DR WPI: 2002-691654/74.

DR N-PSDB; AAD46278.

XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
XX of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
XX analyte detection assays or fluorescence activated cell sorting  
XX applications.

PS Disclosure; Page 70-71; 80pp; English.

XX The invention relates to nucleic acid molecules encoding non-aggregating  
 CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
 CC useful in analyte detection assays, as colouring agents, as markers in  
 CC recombinant DNA applications, as screens or filters, in fluorescence  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, in screening assays, as second  
 CC messenger detectors, in fluorescence activated cell sorting applications,  
 CC in protease cleavage assays or as fluorescent timers. The present  
 CC sequence is Discosoma sp. drfp583 (NFP-6) wild-type protein of the  
 CC invention

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 5; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVTKERFRRVMEGTNGHEPEIEGEGRPYEGHNTYKLTQKGGPLFPAMD 60  
 DB 1 MRSSKNVTKERFRRVMEGTNGHEPEIEGEGRPYEGHNTYKLTQKGGPLFPAMD 60  
 QY 61 LSPQFOYGSKYVVKHPADIPYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120  
 DB 61 LSPQFOYGSKYVVKHPADIPYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120  
 QY 121 KYKFIGVNFPSDGPVMOCKTGWGEASTERTLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
 DB 121 KYKFIGVNFPSDGPVMOCKTGWGEASTERTLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
 QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTVQYERTTGRHHLFL 225  
 DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTVQYERTTGRHHLFL 225

RESULT 6  
 AAE17540 standard; protein; 225 AA.

AC AAE17540;  
 DT 22-APR-2002 (first entry)  
 DE Discosoma sp. humanised wild-type Anthozoa protein drfp583.  
 XX Fluorescent timer protein; protein movement; translocation; trafficking;  
 KW promoter activity; gene expression; transgenic plant; gene modification;  
 KW protein age; anthozoa protein; drfp583.

OS Discosoma sp.  
 XX WO200196373-A2.  
 PN 20-DEC-2001.  
 PD 13-JUN-2001; 2001WO-US019097.  
 PF 14-JUN-2000; 2000US-0211607P.  
 PR (CLON-) CLONTECH LAB INC.  
 PA Pradkov AF, Terékikh A;  
 PI WPI; 2002-154595/20.  
 DR N-PSDB; AAD28207.  
 XX New fluorescent timer proteins comprising an emission spectrum that  
 PT changes over time from a first wavelength to a second wavelength, useful  
 PT for monitoring intracellular protein movement, translocation, trafficking  
 PT or stability.  
 PR Example 1; Fig 1; 89pp; English.

XX The invention relates to a fluorescent timer protein having an emission  
 CC spectrum that changes over time after synthesis from a first wavelength  
 CC to a second wavelength. The fluorescent timer proteins are useful in  
 CC monitoring the activity of a promoter, determining the age of a protein,  
 CC identifying an agent that modulates the activity of a promoter and in  
 CC enriching a population of cells comprising a fluorescent timer protein.  
 CC The fluorescent timer proteins are also useful for assessing gene  
 CC expression during development of a multicellular organism or during  
 CC cellular differentiation, in response to a drug or other inducer of  
 CC promoter activity, as a reporter to serve as a read-out of promoter  
 CC activity, monitoring intracellular protein movement or translocation,  
 CC protein trafficking, or protein stability, to investigate temporal  
 CC aspects of the activity of a regulatory element, for determining cell  
 CC fate during development and organ remodeling, in spatial and temporal  
 CC visualisation of newly synthesised proteins and accumulated proteins, and  
 CC in distinguishing between newly formed and pre-existing structures, e.g.  
 CC membrane junctions and extracellular matrix components. The fluorescent  
 CC timer proteins may further be used to investigate where photobleaching  
 CC techniques are employed, as detectable labels, as selectable markers, as  
 CC biosensors in prokaryotic and eukaryotic cells, in protease cleavage  
 CC assays, and as second messenger detectors. The nucleic acids can be used  
 CC to generate transgenic, non-human plants or animals or site-specific gene  
 CC modifications in cell lines. The present sequence is Discosoma sp.  
 CC humanised wild-type Anthozoa protein drfp583 used for generating  
 CC fluorescent proteins

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 5; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVTKERFRRVMEGTNGHEPEIEGEGRPYEGHNTYKLTQKGGPLFPAMD 60  
 DB 1 MRSSKNVTKERFRRVMEGTNGHEPEIEGEGRPYEGHNTYKLTQKGGPLFPAMD 60  
 QY 61 LSPQFOYGSKYVVKHPADIPYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120  
 DB 61 LSPQFOYGSKYVVKHPADIPYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120  
 QY 121 KYKFIGVNFPSDGPVMOCKTGWGEASTERTLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
 DB 121 KYKFIGVNFPSDGPVMOCKTGWGEASTERTLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
 QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTVQYERTTGRHHLFL 225  
 DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTVQYERTTGRHHLFL 225

RESULT 7  
 AAO18270 standard; protein; 225 AA.

AC AAO18270;  
 DT 26-SEP-2002 (first entry)  
 DE Discosoma red fluorescent protein.  
 XX Yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette;  
 KW modified yeast strain; environmental pollution.

OS Discosoma sp.  
 XX DE10061872-A1.  
 PN 20-JUN-2002.  
 PD 12-DEC-2000; 2000DE-01061872.  
 PR 12-DEC-2000; 2000DE-01061872.

PA (LICH/) LICHTENBERG-FRATE H.  
 XX  
 PI Lichtenberg-Frate H;  
 XX  
 DR WPI; 2002-539633/58.  
 XX  
 DR N-PSDB; AAL47952.  
 XX  
 PT Modified yeast strain, useful for detecting toxic compounds in  
 PT environment, contains integrated cassettes responsive to genotoxic and  
 PT cytotoxic compounds.  
 XX  
 PS Disclosure, page 21-22; 34pp; German.  
 XX  
 CC The present invention relates to a modified yeast strain that contains,  
 CC integrated stably and functionally in its genome, a genotoxicity cassette  
 CC and a cytotoxicity cassette, each comprising a promoter and reporter  
 CC gene, both of which are different in the two cassettes. The modified  
 CC yeast strain is used to detect environmental pollution, especially  
 CC genotoxic and/or cytotoxic substances in complex environmental  
 CC contaminants, especially organic compounds, but also (non-)ionising  
 CC radiation and chemical carcinogens. Particular applications are in  
 CC monitoring (waste) water (e.g. as an early warning system), medical  
 CC toxicology screening and for industrial process control. The present  
 CC sequence is a marker protein suitable for use in the cassettes of the  
 CC present invention.  
 XX  
 SQ Sequence 225 AA;  
 Query Match 100.0%; Score 1214; DB 5; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSSSKVVIKEFMFKYRMGTVNGHFEIIEGEGEGRPYEGSHNTVTKLVTKGGLPPAMD 60  
 DB 1 MSSSKVVIKEFMFKYRMGTVNGHFEIIEGEGEGRPYEGSHNTVTKLVTKGGLPPAMD 60  
 QY 61 LSPQFYGSKVYVKHPADIPDYKLSPPGFKMERVMNFDGCVTVTQDSSIQDGCFTY 120  
 DB 61 LSPQFYGSKVYVKHPADIPDYKLSPPGFKMERVMNFDGCVTVTQDSSIQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGVKGEIHKALKKOGSHLVPEKSI 180  
 DB 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGVKGEIHKALKKOGSHLVPEKSI 180  
 QY 181 YMAKKFVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 DB 181 YMAKKFVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 8  
 ADY51716  
 ID ADY51716 standard; protein: 225 AA.  
 XX  
 AC ADY51716;  
 XX  
 DT 05-MAY-2005 (first entry)  
 XX  
 DE Discosoma sp. wild type red fluorescent protein (RFP) Seq 12.  
 XX  
 KW fluorescence; mutagenesis; red fluorescent protein; protein interaction.  
 XX  
 OS Discosoma sp.  
 XX  
 PN W0200268605-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 26-FEB-2002; 2002MO-US006063.  
 XX  
 PR 26-FEB-2001; 2001US-00794308.  
 XX  
 PR 24-MAY-2001; 2001US-00866538.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.

XX  
 PI Tsien RY, Baird GS, Campbell RE, Zacharias DA;  
 XX  
 DR WPI; 2002-713372/77.  
 XX  
 DR N-PSDB; ADY51715.  
 XX  
 PT New non-oligomerizing fluorescent protein containing at least one  
 PT mutation that reduces or eliminates the ability of the protein to  
 PT oligomerize, useful for making better and new assays for molecular  
 PT biology.  
 XX  
 PS Claim 10; SEQ ID NO 12; 117pp; English.  
 XX  
 CC This invention relates to a novel non-oligomerizing fluorescent protein.  
 CC Specifically, it refers to the presence or at least one mutation in the  
 CC fluorescent protein that reduces or eliminates the ability of the protein  
 CC to oligomerize. The present invention describes fluorescent proteins and  
 CC derived from naturally occurring green or red fluorescent proteins and  
 CC provides a fusion protein that comprises a non-oligomerizing fluorescent  
 CC protein linked to at least one protein of interest. As such, these fusion  
 CC proteins can be used in methods and compositions to determine the pH of a  
 CC sample, or whether the sample contains an enzyme, molecule or agent that  
 CC regulates the activity of an expression control sequence. Furthermore,  
 CC they may be used to identify a specific interaction of molecules, such  
 CC that they are useful for improving or developing new assays in the field  
 CC of molecular biology. This polypeptide sequence is the Discosoma sp. wild  
 CC type red fluorescent protein (RFP) of the invention.  
 XX  
 SQ Sequence 225 AA;  
 Query Match 100.0%; Score 1214; DB 5; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSSSKVVIKEFMFKYRMGTVNGHFEIIEGEGEGRPYEGSHNTVTKLVTKGGLPPAMD 60  
 DB 1 MSSSKVVIKEFMFKYRMGTVNGHFEIIEGEGEGRPYEGSHNTVTKLVTKGGLPPAMD 60  
 QY 61 LSPQFYGSKVYVKHPADIPDYKLSPPGFKMERVMNFDGCVTVTQDSSIQDGCFTY 120  
 DB 61 LSPQFYGSKVYVKHPADIPDYKLSPPGFKMERVMNFDGCVTVTQDSSIQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGVKGEIHKALKKOGSHLVPEKSI 180  
 DB 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGVKGEIHKALKKOGSHLVPEKSI 180  
 QY 181 YMAKKFVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 DB 181 YMAKKFVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 9  
 AAE34962  
 ID AAE34962 standard; protein: 225 AA.  
 XX  
 AC AAE34962;  
 XX  
 DT 28-MAY-2003 (first entry)  
 XX  
 DE Discosoma species red fluorescent protein (RFP).  
 XX  
 KW phosphorylation indicator; fluorescent protein; detection; phosphatase;  
 XX  
 OS Discosoma sp.  
 XX  
 PN W0200295058-A2.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 24-MAY-2002; 2002MO-US016955.  
 XX  
 PR 24-MAY-2001; 2001US-00865291.  
 XX

XX	PA	(RGC ) UNIV CALIFORNIA.
XX	PI	Tsien RY, Ting AY, Zhang J;
XX	PI	WPI; 2003-148474/14.
DR	N-PSDB;	AAD53432.
XX	PT	Novel chimeric phosphorylation indicators, useful for detecting
PT	kinase/phosphatase in samples, has donor molecule, phosphorylatable	
PT	domain, phosphoaminoacid binding domain, and acceptor molecule, in	
XX	PT	operative linkage.
XX	PS	Di Disclosure; Col 65-66; 38pp; English.
CC	XX	The present invention relates to chimeric phosphorylation indicators
CC	CC	comprising a phosphorylation polypeptide and a fluorescent protein or in
CC	CC	operative linkage, a donor molecule, a phosphorylatable domain, a
CC	CC	phosphoaminoacid binding domain (PABD) and an acceptor molecule. The
CC	CC	phosphorylation indicators of the invention are useful for detecting
CC	CC	kinases or phosphatases in a biological sample. They are also useful in
CC	CC	high throughput analysis e.g. for detecting a kinase inhibitor or
CC	CC	phosphatase inhibitor. The present sequence is Discosoma species red
CC	CC	fluorescent protein (RFP) used in the invention
XX	SQ	Sequence 225 AA;
Query Match		100.0%; Score 1214; DB 6; Length 225;
Best Local Similarity		100.0%; Pred. No. 1e-127;
Matches 225;	Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	1	MRSKKNVKEFMRFKRVMEGTVNGHEFEIEGEGRGPRPYEHNVTKLKYTKGGPLFPAMD I 60
Db	1	MRSKKNVKEFMRFKRVMEGTVNGHEFEIEGEGRGPRPYEHNVTKLKYTKGGPLFPAMD I 60
QY	61	LSPQFQYSKKYYVHGPADIPYKLSPEEGFKWERVMNPFDCGVTVTVQDSLDGCFIY 120
Db	61	LSPQFQYSKKYYVHGPADIPYKLSPEEGFKWERVMNPFDCGVTVTVQDSLDGCFIY 120
QY	121	KVKRTIGVFPEDGPPMOKKTGMWEASTERYLPDGLKGEIHKALKLKDGHYLVEPKSI 180
Db	121	KVKRTIGVFPEDGPPMOKKTGMWEASTERYLPDGLKGEIHKALKLKDGHYLVEPKSI 180
QY	181	YMAKKPVOLPGYYYVDISKLDITSHNEDTYTVEQYERTGRHHFL 225
Db	181	YMAKKPVOLPGYYYVDISKLDITSHNEDTYTVEQYERTGRHHFL 225
RESULT 10		
ID	ADC24126	standard; protein; 225 AA.
XX	AC	ADC24126;
XX	DT	18-DEC-2003 (first entry)
DE	XX	Discosoma wild-type red fluorescent protein.
KX	KM	Discosoma red fluorescent protein; Dared; AB interface; AC interface;
KM	KM	fluorescent protein variant; transcription induction detection;
KM	KM	fluorescence energy resonance transfer; FRET; protein kinase;
OS	XX	protein phosphatase; ion indicator.
XX	XX	Discosoma.
XX	PN	US2003059835-A1.
XX	PD	27-MAR-2003.
XX	PF	10-APR-2002; 2002US-00121258.
XX	RR	26-FEB-2001; 2001US-00794308.
XX	RR	24-MAY-2001; 2001US-0086538.

Query Match	Best Local Similarity	100.0%	Score 1214	DB 7	Length 225
Matches 225	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	1	MRSSKNVYKEFMRRKRVNMGTVNGHBEIEBEGEGRPYEGHNTYVKLYTKGGLPFPAMD	60		
DB	1	MRSSKNVYKEFMRRKRVNMGTVNGHBEIEBEGEGRPYEGHNTYVKLYTKGGLPFPAMD	60		
QY	61	LSPOFGVSKYVYVHPADIPYKKLSFPEEGKMRVVMFEEGGVTVYQDSSLQDGCFTY	120		
DB	61	LSPOFGVSKYVYVHPADIPYKKLSFPEEGKMRVVMFEEGGVTVYQDSSLQDGCFTY	120		
QY	121	KVKFIGNVFPDSGDVPMQKKTGWGWEASTERTLRPRDGLVKGELHKKALKDQGGHYLVEFNSI	180		
DB	121	KVKFIGNVFPDSGDVPMQKKTGWGWEASTERTLRPRDGLVKGELHKKALKDQGGHYLVEFNSI	180		
QY	181	YMAKKPVQLPGYIYVDSKLDITSHNEDYTYVEQYERTTGRHHLFL 225			
DB	181	YMAKKPVQLPGYIYVDSKLDITSHNEDYTYVEQYERTTGRHHLFL 225			

RESULT 11  
ABM00918  
ID ABM00918 standard; protein; 225 AA.  
XX ABM00918;  
AC

XX 15-JAN-2004 (first entry)  
 DT Discosoma sp. red fluorescent protein (RFP).  
 XX  
 DE Fluorescent protein; resonance energy transfer; pH; detection;  
 KW red fluorescent protein; RFP.  
 XX  
 OS Discosoma sp.  
 XX  
 PN US2003170911-A1.  
 XX  
 PD 11-SEP-2003.  
 XX  
 PF 26-FEB-2001; 2001US-00794308.  
 XX  
 PR 26-FEB-2001; 2001US-00794308.  
 XX  
 PA (TSIE/) TSIE R Y.  
 PA (ZACH/) ZACHARIAS D A.  
 PA (BAIR/) BAIRD G S.  
 XX  
 PI Telen RY, Zacharias DA, Baird GS;  
 PI  
 DR WPI; 2003-802418/75.  
 DR N-PSDB; AAD61969.  
 XX  
 PT Fluorescent proteins containing a mutation that reduces or eliminates its  
 PT ability to oligomerize which gives more reliable fluorescence resonance  
 PT energy transfer results and are useful to detect molecule interaction,  
 PT enzymes, or sample pH.  
 XX  
 PS Claim 10; Page 30-31; Opp: English.  
 XX  
 CC The invention relates to a non-oligomerizing fluorescent protein  
 CC containing a mutation that reduces or eliminates its ability to  
 CC oligomerize. The fluorescent protein gives more reliable fluorescence  
 CC resonance energy transfer results and are useful to detect molecule  
 CC interaction, enzymes, or sample pH. These are also used to identify  
 CC agents or conditions that regulate expression of control sequences. The  
 CC present sequence is Discosoma sp. red fluorescent protein (RFP)  
 CC  
 XX  
 SQ Sequence 225 AA;  
 Query Match 100.0%; Score 1214; DB 7; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127; Indels 0; Gaps 0;  
 Matches 225; Conservative 0; Mismatches 0;  
 QY 1 MRSSKNVKEFMRPKVMBSGTGNGHEFEIBGBEGRPYEGHNTVKLKVTGKGPLPFAMD 60  
 DB 1 MRSSKNVKEFMRPKVMBSGTGNGHEFEIBGBEGRPYEGHNTVKLKVTGKGPLPFAMD 60  
 QY 61 LSPQFOYGSKYVYKHPRADIPDYKKLSFPEGFKMERVNNFEDGGVYVTTQDSSLQDGCFTY 120  
 DB 61 LSPQFOYGSKYVYKHPRADIPDYKKLSFPEGFKMERVNNFEDGGVYVTTQDSSLQDGCFTY 120  
 QY 121 KVFPIGNFSPSDGVMQKKTGMWEASTERYPRDGVYKGEIHKALKKDGSHYIVERKSI 180  
 DB 121 KVFPIGNFSPSDGVMQKKTGMWEASTERYPRDGVYKGEIHKALKKDGSHYIVERKSI 180  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEOYERTEGRRHFL 225  
 DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEOYERTEGRRHFL 225

RESULT 12  
 ID ADF70403 standard; protein; 225 AA.  
 XX ADF70403;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX

DE Discosoma wild-type GFP variant protein Seqid26.  
 XX  
 KW ligand; orphan receptor protein; fusion protein; fluorescent protein;  
 KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
 KW GFPuv; Enhanced GFP; EGFP.  
 XX  
 OS Discosoma sp.  
 XX  
 PN WO2003071272-A1.  
 XX  
 PD 28-AUG-2003.  
 XX  
 PF 21-FEB-2003; 2003WO-JP001901.  
 XX  
 PR 22-FEB-2002; 2002JP-00045728.  
 PR 23-JUL-2002; 2002JP-00213949.  
 PR 11-OCT-2002; 2002JP-00298237.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;  
 PI  
 DR WPI; 2003-697654/66.  
 DR N-PSDB; ADF70404.  
 XX  
 PT Transformation of cells with a fusion protein of an orphan receptor  
 PT protein with a fluorescent protein useful for identification of ligands  
 PT to the orphan receptor.  
 XX  
 PS Disclosure; SEQ ID NO 26; 594pp; Japanese.  
 XX  
 CC This invention relates to a novel method of identifying ligands to an  
 CC orphan receptor protein which comprises transforming cells with DNA  
 CC encoding a fusion protein of the orphan receptor with a fluorescent  
 CC protein, so that the fusion protein is expressed in the cells (or cell  
 CC membranes isolated from them) and contacting the cells with the potential  
 CC ligand to be tested. A suitable fluorescent protein (GFP) for example GFP-1,  
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
 CC identification of ligands binding to an orphan receptor protein.  
 CC  
 XX  
 SQ Sequence 225 AA;  
 Query Match 100.0%; Score 1214; DB 7; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127; Indels 0; Gaps 0;  
 Matches 225; Conservative 0; Mismatches 0;  
 QY 1 MRSSKNVKEFMRPKVMBSGTGNGHEFEIBGBEGRPYEGHNTVKLKVTGKGPLPFAMD 60  
 DB 1 MRSSKNVKEFMRPKVMBSGTGNGHEFEIBGBEGRPYEGHNTVKLKVTGKGPLPFAMD 60  
 QY 61 LSPQFOYGSKYVYKHPRADIPDYKKLSFPEGFKMERVNNFEDGGVYVTTQDSSLQDGCFTY 120  
 DB 61 LSPQFOYGSKYVYKHPRADIPDYKKLSFPEGFKMERVNNFEDGGVYVTTQDSSLQDGCFTY 120  
 QY 121 KVFPIGNFSPSDGVMQKKTGMWEASTERYPRDGVYKGEIHKALKKDGSHYIVERKSI 180  
 DB 121 KVFPIGNFSPSDGVMQKKTGMWEASTERYPRDGVYKGEIHKALKKDGSHYIVERKSI 180  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEOYERTEGRRHFL 225  
 DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEOYERTEGRRHFL 225

RESULT 13  
 ID ADH34489 standard; protein; 225 AA.  
 XX ADH34489;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Discosoma sp. red fluorescent protein Dared (wild-type).  
 XX



XX The invention relates to a polynucleotide sequence encoding a Discosoma  
 CC red fluorescent protein (Dared) variant having a reduced propensity to  
 CC oligomerize. The protein variant comprises one or more amino acid  
 CC substitutions at the AB and/or AC interface(s) of the wild-type Dared  
 CC sequence, where the substitutions result in reduced propensity of the  
 CC Dared variant to form tetramers and where the variant displays detectable  
 CC fluorescence of at least one red wavelength. The composition and methods  
 CC are useful in producing red fluorescent proteins having reduced  
 CC propensity for oligomerization, especially tetramerization. The protein  
 CC may be used in molecular biology and in other scientific applications,  
 CC such as in immunosays or hybridization assays, or in tracking the  
 CC movement of proteins in cells. This sequence corresponds to the Dared  
 CC protein.

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSKNVIEKPMRPFKVRMGVTNGHGFETEGEGGRPYEGHNTVKLKTGKGPLPAMDI 60  
 DB 1 MSSKNVIEKPMRPFKVRMGVTNGHGFETEGEGGRPYEGHNTVKLKTGKGPLPAMDI 60  
 QY 61 LSPQFGYSKVVYKHPADIPDYKCLSPPEGFKMERVMNFEDGGVTVTQDSSLQDCCFTY 120  
 DB 61 LSPQFGYSKVVYKHPADIPDYKCLSPPEGFKMERVMNFEDGGVTVTQDSSLQDCCFTY 120  
 QY 121 KYKFIGVNFPSDGPVNOCKTMGWEASTERLYPRDGYLKGEIHKALKXGGRHLYVEFKSI 180  
 DB 121 KYKFIGVNFPSDGPVNOCKTMGWEASTERLYPRDGYLKGEIHKALKXGGRHLYVEFKSI 180  
 QY 181 YMAKKEVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225  
 DB 181 YMAKKEVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225

RESULT 15

ADN33979 standard; protein; 225 AA.

ADN33979;

01-JUL-2004 (first entry)

wild-type Dared protein.

Childarian; fluorescence resonance energy transfer; FRET; wild-type Dared.

Discosoma sp.

MO2003054158-A2.

03-JUL-2003.

18-DEC-2002; 2002MO-US040539.

19-DEC-2001; 2001US-03411733P.

(UYCH-) UNIV CHICAGO.

Bevis B, Glick B;

WPI, 2003-569236/53.

N-PSDB; ADN33978.

Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent  
 mutant of a Childarian chromo- or fluorescent protein or its mutant,  
 useful for applications involving chromo- or fluorescent proteins.

Claim 8; SEQ ID NO 2; 65bp; English.

CC The present invention relates to nucleic acid that encodes a rapidly  
 CC maturing chromo- or fluorescent mutant of a Childarian chromo- or  
 CC fluorescent protein or its mutant. The protein is useful in applications  
 CC involving nucleic acid encoding a chromo- or fluorescent protein and is  
 CC useful for producing a chromo and/or fluorescent protein which involves  
 CC growing the cell, whereby the protein is expressed, and isolating the  
 CC protein substantially free of other proteins. The protein is useful in  
 CC applications involving chromo- or fluorescent protein and is useful as  
 CC PCR primers, hybridization probes, etc. The expression cassettes are  
 CC useful for synthesizing related proteins. The chromoproteins are useful  
 CC as coloring agents which are capable of imparting color or pigment to a  
 CC particular composition of matter e.g. food compositions, pharmaceuticals,  
 CC cosmetics living organisms, e.g., animals and plants. The chromoproteins  
 CC may also find use as labels in analyte detection assays, e.g. assays for  
 CC DNA applications, e.g. the production of transgenic cells and organisms.  
 CC The fluorescent proteins find use in a variety of different applications,  
 CC e.g. in fluorescence resonance energy transfer (FRET) applications, as  
 CC biosensors in prokaryotic and eukaryotic cells, in applications involving  
 CC the automated screening of arrays of cells expressing fluorescent  
 CC reporting groups by using microscopic imaging and electronic analysis, as  
 CC second messenger detectors, and in fluorescence activated cell sorting  
 CC applications and as in vivo marker in animals. The fluorescent proteins  
 CC also find use in protease cleavage assays. The proteins can also be used  
 CC as assays to determine the phospholipid composition in biological  
 CC membranes and as a fluorescent timer. The present sequence represents the  
 CC wild-type Dared.

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSKNVIEKPMRPFKVRMGVTNGHGFETEGEGGRPYEGHNTVKLKTGKGPLPAMDI 60  
 DB 1 MSSKNVIEKPMRPFKVRMGVTNGHGFETEGEGGRPYEGHNTVKLKTGKGPLPAMDI 60  
 QY 61 LSPQFGYSKVVYKHPADIPDYKCLSPPEGFKMERVMNFEDGGVTVTQDSSLQDCCFTY 120  
 DB 61 LSPQFGYSKVVYKHPADIPDYKCLSPPEGFKMERVMNFEDGGVTVTQDSSLQDCCFTY 120  
 QY 121 KYKFIGVNFPSDGPVNOCKTMGWEASTERLYPRDGYLKGEIHKALKXGGRHLYVEFKSI 180  
 DB 121 KYKFIGVNFPSDGPVNOCKTMGWEASTERLYPRDGYLKGEIHKALKXGGRHLYVEFKSI 180  
 QY 181 YMAKKEVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225  
 DB 181 YMAKKEVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225

Search completed: January 11, 2006, 02:02:58  
 Job time : 136 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 11, 2006, 01:55:51 ; Search time 16 Seconds  
(without alignments)  
1353.048 Million cell updates/sec

Title: US-10-006-922a-12  
Perfect score: 1234  
Sequence: 1 MRSSKNVKEFMRFKRMWG.....EDYIVQYERTGRHFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211.5	17.4	238	1 J01514	green-fluorescent
2	93.5	7.7	26936	1 T38344	tilin, cardiac mus
3	92	7.6	1873	2 T30944	surface protein pr
4	91	7.5	458	1 A42386	hsp 90-binding pro
5	89	7.3	1433	1 A36734	basillopeptidase F
6	87.5	7.2	1116	2 B70476	hypothetical prote
7	87	7.2	568	2 T06489	probable peptidylp
8	87	7.2	1484	2 C97196	probable membrane
9	85.5	7.0	456	2 G69397	signal-transducing
10	85	7.0	271	2 F69442	hypothetical prote
11	84.5	7.0	679	2 A40351	adhesion-type prot
12	84.5	7.0	725	1 IUMSNG	neural cell adhesi
13	84	6.9	1616	2 T17884	8-layer protein -
14	83.5	6.9	268	2 E90276	conserved hypotet
15	83.5	6.9	340	2 E69544	hypothetical prote
16	83.5	6.9	374	2 T06244	gibberellin 3 beta
17	83.5	6.8	374	2 T06244	gibberellin 3 beta
18	83	6.8	373	2 T50605	hypothetical prote
19	83	6.8	15281	2 S41309	cyclosporin synthe
20	82.5	6.8	559	1 S55393	peptidylprolyl iso
21	82.5	6.8	680	2 S17962	Kallmann syndrome
22	82.5	6.8	931	2 T32919	hypothetical prote
23	82	6.8	341	2 B53125	restriction enzyme
24	81.5	6.7	292	2 C69106	glucose-1-phosphat
25	81.5	6.7	296	2 T37989	La 4.1 protein - h
26	81.5	6.7	551	1 S72485	peptidylprolyl iso
27	81.5	6.7	629	2 C64180	hypothetical prote
28	81.5	6.7	6805	2 S20901	ctlin - rabbit (Er
29	81	6.7	346	2 S77025	ntriase (EC 3.5.

30	81	6.7	587	1 E69171	phosphoesterase-re
31	80.5	6.6	404	1 S03849	ribonucleoprotein
32	80.5	6.6	415	1 UCI494	ribonucleoprotein
33	80.5	6.6	862	2 F75116	hypothetical prote
34	80.5	6.6	1003	1 AJCHPR	phosphoribosylamin
35	80.5	6.6	1115	1 IUMSNG	neural cell adhesi
36	80	6.6	621	2 A95250	cholesterol-binding pr
37	80	6.6	690	2 P98114	cholesterol-binding pr
38	79.5	6.5	459	2 A46372	immunophilin FKBP5
39	79.5	6.5	475	2 T44566	conserved hypotet
40	79.5	6.5	487	2 T45982	hypothetical prote
41	79.5	6.5	648	2 B84139	ABC transporter (p
42	79.5	6.5	853	1 IJBONC	neural cell adhesi
43	79	6.5	458	2 UN0873	immunophilin p59 -
44	79	6.5	703	2 A64351	hypothetical prote
45	79	6.5	1214	2 T47438	disease resistance

ALIGNMENTS

RESULT 1  
J01514  
green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)  
C/Species: Aequorea victoria  
C/Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 31-Dec-2004  
C/Accession: J01514; P00335; S48693; S51330; S51331  
R/Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.  
Gene 111, 229-233, 1992  
A/Title: Primary structure of the Aequorea victoria green-fluorescent protein.  
A/Reference number: J01514; PMID:92175527; PMID:1347277  
A/Accession: J01514  
A/Molecule type: DNA  
A/Residues: 1-107, 'S', 109-238 <PRA1>  
A/Cross-references: UNIPROT:P42212; UNIPROT:Q17106; UNIPROT:Q17105; UNIPARC:UPI000016B7P4  
A/Residues: 1-107, 'S', 109-238 <PRA1>  
A/Molecule type: DNA  
A/Residues: 1-99, 'P', 101-140, 'L', 142-218, 'V', 220-238 <PRA2>  
A/Cross-references: UNIPARC:UPI0000002FB4; GB:M62853; NID:9155660; PIDN:AAA27721.1; PID:5  
A/Accession: P00335  
A/Molecule type: protein  
A/Residues: 46-64; 74-122; 132-151; 154-183; 185-200 <PRA3>  
A/Cross-references: UNIPARC:UPI0000173FD7; UNIPARC:UPI0000173FD8;  
R/Inouye, S.; Tsuji, F.I.  
FEBS Lett. 351, 211-214, 1994  
A/Title: Evidence for redox forms of the Aequorea green fluorescent protein.  
A/Reference number: S48693; PMID:94364470; PMID:8082767  
A/Accession: S48693  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-24, 'Q', 26-156, 'P', 158-171, 'K', 173-238 <INO>  
A/Cross-references: UNIPARC:UPI000003582F; GB:129345; NID:9606383; PIDN:AAA58246.1; PID:5  
R/Watkins, J.N.; Campbell, A.K.  
Submitted to the EMBL data library, January 1995  
A/Reference number: S51330  
A/Accession: S51330  
A/Molecule type: mRNA  
A/Residues: 1-13, 'V', 15-24, 'Q', 26-44, 'N', 46-153, 'G', 155-156, 'P', 158-171, 'K', 173-227, 'R', 2  
A/Cross-references: UNIPARC:UPI0000079E2D; EMBL:X83959; NID:9634008; PIDN:CAA58789.1; PII  
A/Experimental source: clone gfp1  
A/Accession: S51331  
A/Molecule type: mRNA  
A/Residues: 1-24, 'Q', 26-29, 'R', 31-83, 'L', 85-153, 'G', 155-156, 'P', 158-171, 'K', 173-208, 'Q', 2  
A/Cross-references: UNIPARC:UPI0000079E2D; EMBL:X83960; NID:9634010; PIDN:CAA58790.1; PII  
R/Yang, F.; Moss, L.G.; Phillips Jr., G.N.  
Submitted to the Brookhaven Protein Data Bank, August 1996  
A/Reference number: A65692; PDB:1GFL  
A/Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A', 2-79, 'R', 81-95  
R/Yang, F.; Moss, L.G.; Phillips Jr., G.N.  
Nat. Biotechnol. 14, 1246-1251, 1996  
A/Title: The molecular structure of green fluorescent protein.

A:Reference number: A58953; MUID:98294543; PMID:9631087  
A:Contents: annotation; X-ray crystallography, 1.9 angstroms  
A:Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJFNV) emitting  
C:Comment: The chromophore of this protein is formed by modification of Ser-dehydro-Tyr-  
A:Gene: GFP  
A:Gene: GFP  
A:Introns: 69/3; 167/3  
C:Keywords: chromoprotein; luminescence  
P:65-67/Cross-link: 5-Imidazopyrrolone (Ser-Gly) #status experimental  
P:66/Modified site: dehydrocytosine (Tyr) #status experimental

Query Match 17.4%; Score 211.5; DB 1; Length 238;  
Best Local Similarity 25.7%; Pred. No. 3.3e-11;  
Matches 56; Conservative 47; Mismatches 102; Indels 13; Gaps 6;

QY 16 VMMEGVNHEFEIIEBEGRGREYEGHNTYKAKVGPPLPFAMDLISPOFOGSKVYVH 75  
DB 16 VALDDGVNCHKSVSESGEGDATTYKLTILKFTCTT-GKLPVMPPTLVTTFSYGVCFSKY 74  
QY 76 PADIP--DYKKLSPEGFKMERVMNPFEDGAVTVTQDSLQDCFTYKVFIGNVPPSDG 133  
DB 75 PPHMKGHPFKSAMPBGVQERTIFPKDGNKTAFAVFPEDTLVNRLELKGIDPFKEHG 134  
QY 134 PVMQKKTGMWEASTERLYPRDGLKGEIHKALKLK---DGGHYLVER--KSIYAKKPV 187  
DB 135 NILGKH-MEYNNVSHVYIMADKQKNGIKVNFKIRINIEDSGVQLADHYQVTPIGDGPV 193  
QY 188 QLPGYVYVDSKLDIT---SHNDYTVVEQYERTEGHH 222  
DB 194 ILPDNHYLSTQSALSKDPNEKRDHMLLEFVTAGITTH 231

RESULT 2  
138344  
titin, cardiac muscle [validated] - human  
N:Alternate names: connectin  
M:Contains: serine/threonine-specific protein kinase (BC 2.7.1.-)  
C:Species: Homo sapiens (man)  
C:Date: 12-Aug-1996 #sequence revision 12-Aug-1996 #ext change 09-Jul-2004  
A:Accession: I38344; I38345; S20898; S20899; S63655; S37393  
R:Label: S.; Koilmeier, B.  
Science 270, 293-296, 1995  
A:Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.  
A:Reference number: A57430; MUID:96026330; PMID:7569978  
A:Accession: I38344  
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EME  
A:Molecule type: mRNA  
A:Residues: 1-26926 <LAB1>  
A:Cross-references: UNIPROT:Q10466; UNIPARC:UPI000017CF2C; EMBL:X90568; NID:g1017424; PI  
R:Musco, G.; Tzatzios, C.; Schuck, P.; Pastore, A.  
Biochemistry 34, 553-561, 1995  
A:Title: Dissecting titin into its structural motifs: identification of an alpha-helix in  
A:Reference number: I38345; MUID:95119041; PMID:7819249  
A:Accession: I38345  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1977-2014 <MUS>  
A:Cross-references: UNIPARC:UPI000016B0E9; EMBL:X83270; NID:g602579; PIDN:CAA58243.1; PI  
R:Label: S.; Gantel, M.; Lakey, A.; Tyndick, J.  
EMBO J. 11, 1711-1716, 1992  
A:Title: Towards a molecular understanding of titin.  
A:Reference number: S20897; MUID:92258380; PMID:1582406  
A:Accession: S20898  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 13597-14200, '1', 14202-14696 <LAB2>  
A:Cross-references: UNIPARC:UPI000016B0E7; EMBL:X64698; NID:g37192; PIDN:CAA45939.1; PID  
A:Accession: S20897  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 16310-16382, 'S', 16384-16756, 'P', 16758-16860 <LAB3>  
A:Cross-references: UNIPARC:UPI000016B0E6; EMBL:X64699; NID:g37190; PIDN:CAA45940.1; PID

A:Accession: S20899  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 'P', 22278-22431, 'R', 22433-22448, 'G', 22450-22453, 'Q', 22455-22480, 'TR', 22483-22  
A:Cross-references: UNIPARC:UPI000017571; EMBL:X64697; NID:g37190; PIDN:CAA45938.1; PID  
R:Koilmeier, B.; Oliveri, N.; Wilt, C.C.; Herrmann, B.G.; Labelt, S.  
J. Mol. Biol. 256, 556-563, 1996  
A:Title: Genomic organization of M line titin and its tissue-specific expression in two c  
A:Reference number: S63655; MUID:96177761; PMID:8604138  
A:Accession: S63655  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 26729-26825 <KOL>  
A:Cross-references: UNIPARC:UPI0000172572; EMBL:X92412; NID:g1236761  
R:Gantel, M.; Leonard, K.; Labelt, S.  
EMBO J. 12, 3827-3834, 1993  
A:Title: Phosphorylation of KSP motifs in the C-terminal region of titin in differentiat  
A:Reference number: S37393; MUID:94008990; PMID:8404852  
A:Accession: S37393  
A:Molecule type: mRNA  
A:Residues: 26831-26926 <GAU>  
A:Cross-references: UNIPARC:UPI0000172573  
R:Improta, S.; Politou, A.S.; Pastore, A.  
Submitted to the Brookhaven Protein Data Bank, February 1996  
A:Reference number: A66736; PDB:1TIT  
A:Contents: annotation; conformation by (1)H-NMR, residues 5253-5341  
R:Fiuh, M.; Pastore, A.  
Submitted to the Brookhaven Protein Data Bank, August 1996  
A:Reference number: A66201; PDB:1NCT  
A:Contents: annotation; conformation by (1)H-NMR, residues 'S', 26059-26155  
C:Genetics:  
A:Gene: GDB:TTN  
A:Cross-references: GDB:127867; OMIM:188840  
A:Map position: 2q31-2q32  
C:Function:  
A:Description: structural protein forming filaments in striated muscle  
C:Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; prot  
C:Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; glyco  
structural protein  
P:24752-35008/Domain: protein kinase homology <RIN>  
P:84,177,905,2276,2378,2459,2481,2563,2669,2763,2896,3088,3179,3384,3432,3628,3772,4068,4  
96,11066,11488,11635,11949,12170,12478,12526,12645,12875,13001,13036,13295,13540,1  
tetus predicted  
P:16780,16976,17579,17602,17667,17681,17845,17899,18121,18188,18209,18336,18670,18680,18  
21900,21935,22295,22495,22627,22897,23024,23318,23883,24012,24177,24290,24447,24642,2484  
P:26171,26178,26184,26190/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 7.7%; Score 93.5; DB 1; Length 26926;  
Best Local Similarity 20.7%; Pred. No. 4.2e+02;  
Matches 51; Conservative 34; Mismatches 84; Indels 77; Gaps 13;

QY 15 KYRHEGYVNGHEFF-----IEBEGRGREYEGHNTV-----KLKTYKGSP, 54  
DB 23493 RYRSTGLTEGLYEHRYVTAIARSGSRPSKPIVAMPDPAPEGKQPMPVTTDTRTSV 23552  
QY 55 PFAMDIISPOFOYSKV--YV-----KH-----PADIPYKKLSPEGFKE--RV 96  
DB 23553 SLANSV--PEBEGSKTYGVLIEMQKVDQHEWTKCNTPTKIRRETYTLHLQGAERYRRV 23610  
QY 97 MNFEDG-----GVVTVTQ-----DSSLQDCFTYKVFIGNVPPSDG--PV 135  
DB 23611 LACNAGGPGPEPAVGVTKVTEMLEYPDYEDERYQEGIFVQGGVIRLRTPIKGPPEPI 23670  
QY 136 MQKTTGMFASTELLYPRDGLKGEIHKALKLK---DGGHYLVER-----KSIYAK 184  
DB 23671 CK-----WTKGGODISKRAMTATSEHTELVIKADRGDSGTVDLVENKCGKAAVYIKV 23725  
QY 185 KPVQLP 190  
DB 23726 RVIGSP 23731

RESULT 3

T30944  
surface protein precursor - Enterococcus faecalis  
C|Species: Enterococcus faecalis  
C|Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C|Accession: T30944  
R|Shankar, V.; Baghdayan, A.S.; Hnycke, M.M.; Lindahl, G.; Gilmore, M.S.  
Infect. Immun. 67, 193-200, 1999  
A|Title: Infection-derived Enterococcus faecalis strains are enriched in esp, a gene encoding a surface protein precursor  
A|Status: preliminary; translated from GB/EMBL/DBJ  
A|Molecule type: DNA  
A|Residues: 1-1873 <SHA>  
A|Cross-references: UNIPROT:Q924N7, UNIPARC:UPI00000B2A27, EMBL:AF034779, NID:G3873186;

Query Match 7.6%; Score 92; DB 2; Length 1873;  
Best Local Similarity 20.8%; Pred. No. 17;  
Matches 46; Conservative 33; Mismatches 56; Indels 86; Gaps 13;

Qy 30 EGESEGRPYEHNIVKLVKTKGGLPRAMDLSPOFGS-----KYVYHNP 77  
Db 708 KGVSESEPIIG-TIPIKIV-----YDQSGTTDLAVTSKNIYENPGE 750  
Qy 78 DIP-DYKLSPEEGPKMERVWNPEDGIVVTQDSSLDGCFIYKVFIGNPSPDS-PV 135  
Db 751 NIPGKHKVITTAG-----EGTSIESGTTVAVK-DGSLPEDKLPV 791  
Qy 136 MOKKTMGWEASTERLYPRDGV-LKGE---IHAKLKIKD-----GHHYLVFP- 177  
Db 792 LKAK---DGVTDKMKPEATQPIKADPTFVSSATKLDIIENPGDNIIPAGYHKVTFETA 847  
Qy 178 -----KSIVAKKRVOLPGYIYVDSKLDITSHEDYT 209  
Db 848 GEGTSIESGTTVFAVKGVSLP-----EDKLPVLKAKDGYT 883

RESULT 4  
A42386  
hsp 90-binding protein p59 - rabbit  
C|Species: Oryctolagus cuniculus (domestic rabbit)  
C|Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C|Accession: A42386  
R|Lebeau, M.C.; Messol, N.; Herrick, J.; Faber, L.E.; Renoir, J.M.; Radanyi, C.; Baulieu  
J. Biol. Chem. 267, 4281-4284, 1992  
A|Title: p59, an hsp 90-binding protein. Cloning and sequencing of its cDNA and preparat  
A|Reference number: A42386; MUID:92165768; PMID:1537818  
A|Accession: A42386  
A|Status: preliminary; not compared with conceptual translation  
A|Molecule type: mRNA  
A|Cross-references: UNIPROT:P27124, UNIPARC:UPI0000148548; GB:M84988; NID:G165594; PIDN  
A|Experimental source: liver  
A|Note: sequence extracted from NCBI backbone (NCBI:83839)  
C|Superfamily: human FK506-binding protein FKBP5; BKR-type peptidylprolyl isomerase h  
F/50-97/Domin: BKR-type peptidylprolyl isomerase homology <PT1>  
F/319-352/Domin: tetratricopeptide repeat homology <TT1>  
F/353-366/Domin: tetratricopeptide repeat homology <TT2>

Query Match 7.5%; Score 91; DB 1; Length 458;  
Best Local Similarity 24.6%; Pred. No. 31;  
Matches 42; Conservative 26; Mismatches 49; Indels 54; Gaps 10;

Qy 57 AMDI-----LSPOFGSKVYVGHPIADIPDYKLSPE--GFKMERVWNP 99  
Db 89 AMDIADVATMKVGLGRITCKPEYAYGS---AGSPKIPPNATLVFEVLPFKGSDLDLD 145  
Qy 100 EDGIVVVTQDSSLDGCFIYKVFIGNV--PSDGPVMOKKTMGWEASTERLYP----R 153  
Db 146 EDGGI-----IRRIITRGEAYAPNDGAIETVALEGY--YKRLFPQREIR 189  
Qy 154 DGVAKGE-----IHAKL-KLKGHHYLVFKSYT---MAKKVVOVLPGY 192  
Db 190 FEVVGESGLDPCGLEKAIQRMEKEHSHLYLKPSAFGNAGKEKFOIPIY 240

RESULT 5

Bacillopeptidase F (BG 3.4.21.-) precursor bpr [validated] - Bacillus subtilis

A:66734

C:Species: Bacillus subtilis

C:Date: 19-May-2000 #sequence revision 19-May-2000 #text change 09-Jul-2004

C:Accession: A36734, A35131, A35750, B35750, S08823, JN0353, I39849, B69596, JN0084

R:Stroma, A.; Rifo Jr., G.A.; Rudolph, C.F.; Sullivan, B.J.; Theriault, K.A.; Pero, J.

J: Bacillus. 172, 5520-5521, 1990

A:Reference number: A36734; MUID:90368623; PMID:2118514

A:Contents: extratum

A:Accession: A36734

A:Molecule type: DNA

A:Residues: 1-1433 <SLQ>

A:Cross-references: UNIPROT:P16397, UNIPARC:UPI000060385, GB:M29035, NID:g143307, PIDN:g143307, PIDN:g143307

R:Stroma, A.; Rifo Jr., G.A.; Rudolph, C.F.; Sullivan, B.J.; Theriault, K.A.; Pero, J.

J: Bacillus. 172, 1470-1477, 1990

A:Title: Bacillopeptidase F of Bacillus subtilis: purification of the protein and cloning

A:Reference number: A35131; MUID:90170864; PMID:2106512

A:Accession: A35131

A:Molecule type: DNA

A:Residues: 1-365, 'S', 367-682, 'EIMP', 893, 'Q', 895-896 <SL2>

A:Cross-references: UNIPARC:UPI0000172C26; GB:M29035

A:Note: the authors translated the codon GAA for residue 545 as Leu

R.Wu, X.C.; Nathoo, S.; Pang, A.S.H.; Carne, T.; Wong, S.L.

J: Biol. Chem. 265, 6845-6850, 1990

A:Title: Cloning, genetic organization, and characterization of a structural gene encoding

A:Reference number: A35750; MUID:90216713; PMID:2108961

A:Accession: A35750

A:Molecule type: DNA

A:Residues: 1-392, 'V', 394-828, 'NIRTLVSLKFRSHKSV', <WUA>

A:Cross-references: UNIPARC:UPI000016E7BA; GB:J05400; NID:g142607; PIDN:AAA83362.1; PID:g142607

A:Note: this sequence has been corrected

A:Accession: B35750

A:Molecule type: DNA

A:Residues: 876-935, 'CG', <WU2>

A:Cross-references: UNIPARC:UPI00000B1ICF; GB:J05400; NID:g142607; PIDN:AAA83363.1; PID:g142607

A:Note: this sequence has been corrected

R.Maunda, E.S.; Naguchi, H.; Sato, T.; Takeuchi, M.; Kobayashi, Y.

A:Title: Nucleotide sequence of the sporulation gene spoIIIGA from Bacillus subtilis.

A:Reference number: S08823; MUID:90174995; PMID:2106671

A:Accession: S08823

A:Molecule type: DNA

A:Residues: 1410-1433 <MAS>

A:Cross-references: UNIPARC:UPI000016B97A; EMBL:X17344; NID:g40165; PIDN:CAA5224.1; PID:g40165

A:Note: source of this material was Bacillus subtilis (natto)

R.Kato, T.; Yamagata, Y.; Arai, T.; Ichishima, E.

Biosci. Biotechnol. Biochem. 56, 1166-1168, 1992

A:Title: Purification of a new extracellular 90-kDa serine proteinase with isoelectric pI

A:Reference number: JN0353; MUID:93005071; PMID:1368833

A:Accession: JN0353

A:Molecule type: protein

A:Residues: 195-218, 'A' <KAT>

A:Cross-references: UNIPARC:UPI00000B2A24

A:Note: source of this material was Bacillus subtilis (natto)

R.Beall, B.; Lowe, M.; Lutkenhaus, J.

J: Bacteriol. 170, 4855-4864, 1988

A:Title: Cloning and characterization of Bacillus subtilis homologs of Escherichia coli

A:Reference number: I39849; MUID:89008108; PMID:3139638

A:Accession: I39849

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-211 <RES>

A:Cross-references: UNIPARC:UPI000016B868; GB:M28230; NID:g142938; PIDN:AAA2458.1; PID:g142938

R.Kumet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Allion, G.; Azevedo, V.; Beteret, N.M.; Choi, C.; Bron, S.; Brouillette, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, C.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, C.; Fader, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fittz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallazzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koestler, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinot, J.; Laubert, J.; Laubert, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maunda, S.; Maueel, A.;

Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetellé, R.; Rieger, M.; Rivaletta, C.; Rooba, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authier, Schlicht, S.; Schroeder, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Seton, A.; Schuch, M.; Tanaka, A.; Tanaka, T.; Terpe, T.; Tognoni, A.; Tognoni, V.; Uchiyama, T.; Walters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.; Authier, Yoshikawa, H.F.; Zamec, H.; Zamec, H.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A:Reference number: A69580; MUID:198044033; PMID:9384377  
A:Accession: B69596  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1433 <KUN>  
A:Cross-references: UNIPARC:UP10000060385; GB:Z99111; GB:Z99112; GB:AL009126; NID:G26339  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: bpr, bpr  
A:Map position: 135 (degrees)  
C:Superfamily: bacillopeptidase F; subtilisin homology  
C:Keywords: extracellular protein; hydrolase; serine protease  
F:1-10/Domain: signal sequence #status predicted <SIG>  
F:31-194/Domain: propeptide #status predicted <PRO>  
F:195-1433/Product: bacillopeptidase F #status experimental <MAT>  
F:218-466/Domain: subtilisin homology <SBT>  
F:227,274,453/Active site: Asp, His, Ser #status predicted

Query Match  
Best Local Similarity 7.3%; Score 89; DB 1; Length 1433;  
Matches 56; Conservative 26; Mismatches 71; Indels 80; Gaps 14;

QY 17 RMETVNGHEPIEGEGEGRPGYEGHNTVQLKVTGSGPLPFA----- 57  
DB 521 KAEQVS-----VEGDGDRPYYQHKKV-TEAYEGSLPLTLTADNVSVTSKSYKLD 574  
QY 58 ---WILSPQFGYSGKYVVKHPADIPDYK--KLSPEEGKMERVWNFEDG-----VYT 106  
DB 575 QGEWTEITAKRISGHLKGTQARIPDIGNTKLS---YKW---MIHFGGHVSSDVVD 627  
QY 107 VHQDSLDQGCFTYVKF-----IGVN-----PSPDGPVWQKKTGMGEASTERY 151  
DB 628 VVVKPSITNG---YKQDEFTAGGVVASTNNMKEGVSTDP-----NTAASGEKVV 677  
QY 152 PRDGVLCGE-----IHKALKKDGSHYLVFEFSIYMAKPPVQLPGYVV 195  
DB 678 GGN--LTGNVANSAMNLMVPIKAPDSGLFLQFMSMNLEDDPDY-GYVTV 727

RESULT 6  
B70476  
hypothetical protein aq\_2054 - *Aquifex aeolicus*  
C:Species: *Aquifex aeolicus*  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C:Accession: B70476  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O.V.  
A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: B70476  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1116 <AGF>  
A:Cross-references: UNIPROT:O67838; UNIPARC:UP100000567A2; GB:AE000770; NID:G2984274; PI  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: aq\_2054

Query Match  
Best Local Similarity 7.2%; Score 87.5; DB 2; Length 1116;  
Matches 56; Conservative 35; Mismatches 103; Indels 61; Gaps 9;

QY 5 KNVIFKFKVMEGVNGHEPIEGEGE-----GRPEGHNTVQLKVTGSG 52  
DB 164 KDLIDSEYFAKTGGEIKRNTGEILAEVKEIKRNTFTLSGTAKNKAGTINLPVLIN 223

QY 53 PLRFAMDLSPQFGY-----KVVVHPADIPDYKLSF-----PRG----F 91  
DB 224 AKAFKDLIVNKNVNGSIEGIVKNGYLPDKLPLGELVNPBGKIKFTYDVIPEGLLTF 283  
QY 92 KMER-VNPFEDGVVTVQDSLDQGCFTYVKFIGNVFPSPDGPVWQKKTGMGEASTERY 150  
DB 284 SPENLVVDKNTLGNREIRGEFNGKVDFFKMFVKV-----AFTENL 327  
QY 151 YPRDGVLCGEIHKALKKDGSHYLVFEFSIYMAKPPVQLPGYVV--SKLDITSNEDYT 209  
DB 328 EVIIRKFKGVDLFGVNFSGNSINFEFGNSGVAK-----GNLIINKKLEGBRFSFNDP 381  
QY 210 IVEO-----YERTGR 220  
DB 382 VVFQDFNAYLSGEGK 396

RESULT 7  
T06489  
probable peptidylprolyl isomerase (EC 5.2.1.8) FKBP77 - wheat  
C:Species: *Triticum aestivum* (common wheat)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T06489  
R:Brieman, A.  
submitted to the EMBL Data Library, August 1996  
A:Reference number: Z15713  
A:Accession: T06489  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-568 <BRI>  
A:Cross-references: UNIPROT:O04843; UNIPARC:UP100000A075; EMBL:Y07636; PIDN:CAA68913.1  
A:Experimental source: cv. ATIR, 2 day old plants, root tips  
C:Genetics:  
A:Gene: FKBP77  
C:Superfamily: peptidylprolyl isomerase ROP1; BKB-type peptidylprolyl isomerase homology  
C:Keywords: calmodulin binding; cis-trans-isomerase; cyclosporin A binding  
F:54-101/Domain: BKB-type peptidylprolyl isomerase homology <PPI>  
F:165-212/Domain: BKB-type peptidylprolyl isomerase homology <PPI2>  
F:286-334/Domain: BKB-type peptidylprolyl isomerase homology <PPI1>  
F:482-515/Domain: tetrahydropeptide repeat homology <THR>

Query Match  
Best Local Similarity 7.2%; Score 87; DB 2; Length 568;  
Matches 61; Conservative 32; Mismatches 81; Indels 108; Gaps 15;

QY 5 KNVIFK-----FMKFKVMEGT-VNGHEPIEGEGRPGYEGHNTVQLKVTGSGPL 54  
DB 39 KKLVEGEGMDTATATALKVEVHYGTLLDGTGFP--SSRDRGTGP-----FKLEGGQVI 91  
QY 55 PFAND-----ILSPQFGYSGKYVVKHPADIPYKXLSFP----- 88  
DB 92 K-GMDQGIKTKKKGEMASLTTPDLAAGERA---PRTIPNATLRFVLLSWASVQDI 146  
QY 89 -----EGKRVVWNPFGDGVVTVQDSLDQGCFTYVKFIGNVFPSPDGPVW 136  
DB 147 CKDGIKFKVLVEQKME--NPKDLDSEYFAKTGGEIKRNTGEILAEVKEIKRNTFTLSGTAKNKAGTINLPVLIN 192  
QY 137 QKKTGMGEASTERYLPDGVLCGEIHKALK-LKDGSHYLVFEFSIY---MAKPPVQLPG 191  
DB 193 -----EFSV-----KDGFCPLASKAVKTKMKGEKVLTLVKKQYGRFGQGRATVEVG 240  
QY 192 YYYVDSKLDI-----TSHNEDTIV-----EQYER 216  
DB 241 AVPNSTLHIDLQVSWKTLTLIGDDKRIKLVLEGGYER 282

RESULT 8  
G97196  
probable membrane protein [imported] - *Clostridium acetobutylicum*  
C:Species: *Clostridium acetobutylicum*  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: G97196

R.Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J. *Daily, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.*

J. Bacteriol. 183, 4823-4838, 2001

A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C10

A/Reference number: A96900; MUID:21359325; PMID:21359325

A/Accession: C97196

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1484 <KUN>

A/Cross-references: UNIPROT:097664; UNIPARC:UP100000CA4ED; GB:AE001437; PIDN:AAK80358.1;

A/Experimental source: Clostridium acetobutylicum ATCC824

C/Genetics:

A/Genes: CAC2403

Query Match 7.2%; Score 87; DB 2; Length 1484;

Best Local Similarity 20.7%; Pred. No. 33;

Matches 62; Conservative 47; Mismatches 76; Indels 114; Gaps 17;

9 KEFMRPKVMEG-----TVNGHEFEIEG-----EGEGRPYGHNTVKLKVTGCGP 53

766 KNIVPCKTLKKNSFVSENPPEYKLEADIRDKNSNAPYKQIAGDSANMVEVAKGCV 825

54 L-----PAMDI-----LSPQD-----YGSKYVVKHPADIPYKKSF-- 87

826 LKARENPPYKFINISGKNTNKKHLAKFYDANMHEIYSS--YTKPSDSDPESHDFTG 883

88 -----PEGFKMER--VNNFED--GGVTVTVQDSLQDCGFYKVKFGVGFPSDGPVQCK 139

884 ECVSPKTKYMLLEISYKPKPKIKYWKIHDVNIYD-LSEYK-----SNSFPMNKK 934

140 TWGWEASTERLYPRDVLKGEIHKALKDGGHYLV-----EFKSIYM 182

935 VK--KAKTAKVYR-----ALVSKSGGKLVTKVGGKSYVDTPESKSGQFKMIVL 982

183 AKKPVLPQYTVDS-----KLD-----ITSHNEDTYIYQYERTEGRHHLF 224

983 GK-----FHPDSGKNSIKVENKOGFNAINLFTVYIPENYNTVYKLETKADKCNIF 1033

RESULT 9

66397 signal-transducing histidine kinase homolog - Archaeoglobus fulgidus

C/Species: Archaeoglobus fulgidus

C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C/Accession: G69397

R.Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A/Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.

A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea

A/Reference number: A69250; MUID:98049343; PMID:9389475

A/Accession: G69397

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-456 <KLE>

A/Cross-references: UNIPROT:O29083; UNIPARC:UP10000055E0A; GB:AE001022; GB:AE000782; NID

Query Match 7.0%; Score 85.5; DB 2; Length 456;

Best Local Similarity 22.0%; Pred. No. 9.5;

Matches 46; Conservative 25; Mismatches 59; Indels 79; Gaps 10;

26 EFPIEGGERPRYGHNTVTKLTKGKGLPPAMDIISPOQYQSKYVVKHPADIPVKL 85

67 EFPELIG-EGRFYFSR-----YITKDGVRVHWGFTKFTPLGKSYITIGMIDVTKSKEL 120

86 SPFEGFKMERVWVFEDGGVTVTQDS-----SLQDCGFY-----KYKFGVN 128

121 -----EQALNASEEFRTVVDLSLTPVYLLQDGLMIVYVVKAFEBATGYKREIYGRN 172

129 -----FPSDGPVQCK-----TWGWEASTERLYPRDGVTK----- 158

DB 173 PEFLLHPEDRGLVYKRYIEREKGLADTMTWYSM-----RIIRKQGEVWVTARPGRVYR 227

QY 159 -----GEIYHKA-LKAKDGGHYL 174

DB 228 GRPAVAATVVDTEIHKINIELKKKGEYL 256

# RESULT 10

F69442 hypotetical protein AF1543 - Archaeoglobus fulgidus

C/Species: Archaeoglobus fulgidus

C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C/Accession: F69442

R.Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A/Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.

A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea

A/Reference number: A69250; MUID:98049343; PMID:9389475

A/Accession: F69442

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-271 <KLE>

A/Cross-references: UNIPROT:O28729; UNIPARC:UP10000056C85; GB:AE000996; GB:AE000782; NID:

C/Superfamily: Pyrococcus horikoshii hypotetical protein PH1420

Query Match 7.0%; Score 85; DB 2; Length 271;

Best Local Similarity 25.6%; Pred. No. 5.2;

Matches 44; Conservative 18; Mismatches 64; Indels 46; Gaps 9;

64 QFOYSGKVVYKHP-----ADIDPYKK--LSPBEGFKMERVWVFEDGGVY---TVTQDSL 113

65 QTEINSLFYTLNLSIMLGBDLDPDYSFLKRLGLEKAEKRYLSDGVATATYFLQPNAL 124

QY 114 QDCGFYRKFGVGNPPSPGVPWQCKTGMWEASTERLYPRD-----GVLKGEIHKALK 166

DB 125 RDA--YMTSLHLNVRDPEETKLV-----RRYRRETFPGVGYVKKENLEETFY 174

QY 167 LKDGHYLVEFSIY-----MAKKPVQLPGY-----YYVDSKLDITS 203

DB 175 ----ASYIIRDAVIVSFVKSFGSNGFAKQPGVYPLDVTYATSTLSLS 222

RESULT 11

A40351 adhesion-type protein ADMLX - human

C/Species: Homo sapiens (man)

C/Date: 24-Jan-1992 #sequence\_revision 24-Jan-1992 #text\_change 15-Mar-2004

C/Accession: A40351

R.Legouis, R.; Haddad, J.P.; Leveilliers, J.; Claverie, J.M.; Compain, S.; Wunderle, V.; G.; Weisenbach, J.; Petit, C.

Cell 67, 423-435, 1991

A/Title: The candidate gene for the X-linked Kallman syndrome encodes a protein related

A/Reference number: A40351; MUID:92005720; PMID:191827

A/Accession: A40351

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-679 <LEG>

A/Cross-references: UNIPARC:UP1000016B3D2; GB:S60085; NID:9237595; PIDN:AA820108.1; PID:5

Query Match 7.0%; Score 84.5; DB 2; Length 679;

Best Local Similarity 27.0%; Pred. No. 20;

Matches 48; Conservative 20; Mismatches 75; Indels 35; Gaps 10;

46 LKTKGK-----PLPF-----AMDILSPQYQYS---KYVVKHPAD--IDYKGLSPFEG 90

Db 406 VKTRKGGIGIQTOLPFQRRRPTPLBVGAFPYODGOLQVKVYWKTEDPVTYNRYHVMPEBA 465

Qy 91 FKMEVYVNFEDGGVYTVQDSSLDGCFIYKVFIGNVFPSPGVQKTKMGWEASTRL 150

Db 466 CAHNHTTGSSEASSGGTHENYIILQDLSFSCKTKV-----TVQFPRPYS---HSEKAEV 515

Qy 151 Y---PRDGLVKGELIKALK-LKDGHYLVFEKSIYMAKKVQLPGYVYVDSKLDITSH 204

Db 516 FTTTPPCALAKGSKHNPICGCGEAGHYVLSK-----VLAKPENLSASFIVQD-VNITGH 567

RESULT 12

IUMSNG

neural cell adhesion molecule 1 precursor, GPI-anchored splice form - mouse

N:Alternate names: NCAM-120

C:Species: Mus musculus (house mouse)

C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 09-Jul-2004

C:Accession: A29673; S00382; A44290

R:Barthelemy, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chalk, J.C.; Hirsch, M.R.; Fontec

EMBO J. 6, 907-914, 1987

A:Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000

A:Reference number: A29673; MUID:87246524; PMID:3595563

A:Accession: A29673

A:Molecule type: mRNA

A:Residues: 1-725 <BAR>

A:Cross-references: UNIPROT:P13594; UNIPARC:UPI0000028AC3; EMBL:Y00051; NID:953342; PIDN

R:Barbae, J.A.; Chalk, J.C.; Steimeetz, M.; Goridis, C.

EMBO J. 7, 625-632, 1988

A:Title: Differential splicing and alternative polyadenylation generates distinct NCAM v

A:Reference number: S00382; MUID:88283628; PMID:3396534

A:Accession: S00382

A:Molecule type: DNA

A:Residues: 642-656, 'D', 658-725 <BA2>

A:Cross-references: UNIPARC:UPI0000174377; EMBL:X07195

R:Rougon, G.; Marzani, D.R.

J. Biol. Chem. 261, 3396-3401, 1986

A:Title: Structural and immunological characterization of the amino-terminal domain of m

A:Reference number: A44290; MUID:86140120; PMID:3512556

A:Accession: A44290

A:Molecule type: protein

A:Residues: 20-36 <ROU>

A:Cross-references: UNIPARC:UPI0000174376

C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol

C:Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:IUMS

C:Genetics:

A:Gene: NCAM

A:Map position: 9

A:Introns: 701/1

C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu

C:Keywords: alternative splicing; cell adhesion; duplication; heparin binding; membrane

P:1-19/Domain: signal sequence #status predicted <SIG>

P:34-98/Domain: immunoglobulin homology <IMM1>

P:132-191/Domain: immunoglobulin homology <IMM2>

P:152-156/Region: heparin binding #status predicted

P:161-165/Region: heparin binding #status predicted

P:228-290/Domain: immunoglobulin homology <IMM3>

P:263-272/Region: NCAM binding #status predicted

P:323-388/Domain: immunoglobulin homology <IMM4>

P:420-482/Domain: immunoglobulin homology <IMM5>

P:519-596/Domain: fibronectin type III repeat homology <FN3A>

P:625-688/Domain: fibronectin type III repeat homology <FN3B>

P:741-96, 139-189, 235-288, 330-386, 427-480/Duplicate bonds: #status predicted

P:222,316,348,424,450,479/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match 7.0%; Score 84.5; DB 1; Length 725;

Best Local Similarity 24.2%; Pred. No. 21;

Matches 37; Conservative 30; Mismatches 57; Indels 29; Gaps 9;

Qy 91 FKMEVYVNFEDGGVTV---TQDSSLDGCFIYKVFIGNVFPSPGVQKTKMGWEAST 147

Db 548 FTWYDAKANEIGVITIMGLKPEYTS-----RLAALNKGIGELIWPQSESKTPVP 600

Qy 148 ERLYPRDGLKGEIHK-----ALKLKDG-----HYLVFEKSIYMAKKP-VQLP-GYY 193

Db 601 ELASPK---LEGQMGEDNSIKVNLIKQDDGSGPIRHLVYKRALASMKPEIRLPSSGH 657

Qy 194 YVDSK-IDTSHNEDYTVF-QVETREGRHLF 224

Db 658 HVMLKSLDMNAEYVYVAENQGGSKAAHFFV 690

RESULT 13

T17884

S-layer protein - Bacillus circulans

C:Species: Bacillus circulans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T17884

R:Albert-Pivert, E.; Davies, J.

Gene 147, 1-11, 1994

A:Title: Biosynthesis of butirosin in Bacillus circulans NRRL B3112: identification by se

A:Reference number: Z18808; MUID:94374689; PMID:7522196

A:Accession: T17884

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1616 <AUB>

A:Cross-references: UNIPROT:P35824; UNIPARC:UPI0000135A00; EMBL:L20421; NID:9304142; PID:

A:Genetics:

A:Gene: butB

C:Function:

A:Pathway: butirosin biosynthesis

Query Match 6.9%; Score 84; DB 2; Length 1616;

Best Local Similarity 20.6%; Pred. No. 69;

Matches 59; Conservative 48; Mismatches 85; Indels 94; Gaps 17;

Qy 6 NVKEFNR-----FKVMEGTGVNHEFEI---EGSGEGRPYEGHNTVLYKVTYK 50

Db 590 DLQGFIRYSRELGDIVHSFNIRAESIASNEFALDSDHDMERYVYNAADNGQIK--- 646

Qy 51 GGPLEPANDLISPOFQYGSKYVYKHPAD-IDPYKKLSFPEGFKERVNFEDGVV--TV 107

Db 647 -----RLRESAKQGAVALFVNPSNDEVDPQLKTYE-----EVLQNYDVGVVLDRA 692

Qy 108 TQDSSLQDGCFTYK---FIGV-----NRP-----SDGPVMO-----KKT 140

Db 693 RYDNESSADPSDLTVAKFPSPFGANGKQLQNPDDVTVAGVNRKQGPILRDMWFRSKTI 752

Qy 141 MGWEASTRLVPRDGLKGEIHKALKLK-----DGHY-LVEFK---SIYMAK 184

Db 753 KSFTSEVRQLDR---VAAEKGGKIEVSAVYGSWFEIYLVGVHMGSTEFRYDERLRMKD 809

Qy 185 KPVQLPGY---YVDSKLDITSHNEDYTVFQYERT--EGRHHLFL 225

Db 810 KSVYTPGYGESGYV-----KNLDPIMIGAYQTTAPEIEHYITL 847

RESULT 14

E90276

conserved hypothetical protein [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C:Accession: E90276

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyez, M.J.; Chan-Y

Jong, I.; Jeffries, A.C.; Kozera, C.D.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A89139

A:Accession: E90276

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-268 <KUR>

A:Cross-references: UNIPROT:Q97YT6; UNIPARC:UPI00000643B4; GB:AE006641; NID:913814420; P:

C:Genetics:

A:Gene: SSO1221

Query Match	6.9%	Score 83.5;	DB 2;	Length 268;
Best Local Similarity	25.8%;	Pred. No. 7;		
Matches 49;	Conservative 21;	Mismatches 59;	Indels 61;	Gaps 13;

34 EGRPYEGHNTVKLTGK-----PLPFAWDILSPQFYGSKY-----VKHPAD 78

Db 3 EIRPY--YNTFSLPTKSGKSQIVPPPWIIYAI----EMIGVKAYFDPVKVLDDLVPPLLE 56

```

79 IPD-----YKLSFPEGFKMERVMNPEDEGVTVITQSSSLQDGCFIYKVKE---IGVNF 125
      |         |||          |||       ::::   :           |||    |
      |         |||          |||       ::::   :           |||    |
      |         |||          |||       ::::   :           |||    |

```

DB 57 LVDEGEFVI IAKI FIVSGNRWE--MLIBD-----PEBIAKIMEKAIPLAVNINDNFIIF 100

[illegible][illegible]

151 INCESTRY 170

**RESULT 15**

hypothetical protein AF2357 - *Archaeoglobus fulgidus*

C:\Species: Archaeoglobus fulgidus

C:\Date: 05-Dec-1997 #bsequence\_05-Dec-1997 #text\_change 09-Jul-2004

C;ACCESSION: E69544  
P:KJENK H P : CJAY

Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gall, S.; Kirkness, E.F.  
 Glodet, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A.; Authors: Uteerback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; kaime, B.F.; sykes, S. Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Reference number: A69250: MUID:98049343: PMID:93894755

A;Accession: E69544

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;CROSS-references: UNIT

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Query Match	6.9%	Score 83.5;	DB 2;	Length 340;
Best Local Similarity	23.9%	Pred NO	9.6%	

Deletions 29; Duplications 29; Gaps 9  
Matches 38; Mismatches 63; Indels 9

```
Oy |::| : : : : ||::|| :::  
   |::| : : : : ||::|| :::  
41 HNTYKLKVTGGPPLFAMWILSPQOYGSKYYVKHADIPDYKUSFPPEGFWERVMNFE 100
```

Db 144 HAWVEVKINN-----SRVADPTVW---LYVNYPEKPNNNKLMF-INNESWANLIDF- 192

101 DGGVTVTQDSSLQDGCFTKVKKFGVNPSSDGPWQ--KLTMGHEASTBELYPIDGVLK 158

150 CE - TUNISIYATIN KUTUBHANALARIYIN BAYANLARI 1962

251 SPTVNTALATP-----TYKFRIL-----VPTWYFELR 276

Search completed: January 11, 2006, 02:04:50  
Job time : 18 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 11, 2006, 01:54:16 ; Search time 73 Seconds  
(without alignments)

2174.574 Million cell updates/sec

Title: US-10-006-922a-12

Sequence: 1 MRSSKNVKEFWKFRKMEG.....EDYIVEQYERTGRHHLFL 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Uniprot\_05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	*Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	1	RFP DISSP
2	1196	98.5	236	2	Q66ND7_9CNID
3	1193	98.3	225	2	Q6KRF5_discosoma s
4	1191	98.1	236	2	Q66ND8_9CNID
5	1085.5	89.4	230	2	Q9GJ77_9CNID
6	985	81.1	226	2	Q5S3G8_discosoma s
7	972	80.1	226	2	Q5S3G7_9CNID
8	804	66.2	221	2	Q66PW1_ACRTE
9	798	65.7	221	2	Q66PW8_acropora ac
10	797	65.7	221	2	Q66PW0_ACRMT
11	793	65.3	221	1	NFCP_GDYTE
12	793	65.3	221	1	Q66PW9_9CNID
13	782	64.4	221	1	NFCP_MONEP
14	745	61.4	227	2	Q66ND6_DISST
15	729.5	60.1	232	1	GPPL DIEST
16	690	56.8	225	2	Q96JF5_9CNID
17	684	56.3	225	2	Q66ND3_9CNID
18	681	56.1	225	2	Q7Z0W4_9CNID
19	681	56.1	225	2	Q95UA7_9CNID
20	679.5	55.9	227	2	Q7Z0W6_montastraea
21	678.5	55.9	227	2	Q7Z0W8_montastraea
22	678.5	55.9	227	2	Q96J29_9CNID
23	678	55.8	225	2	Q7Z0W5_9CNID
24	677	55.8	224	2	Q5T1L6_echinophylli
25	672.5	55.4	227	2	Q66ND2_9CNID
26	670.5	55.2	227	2	Q66ND5_9CNID
27	669.5	55.1	234	2	Q7Z0W7_9CNID
28	668.5	55.1	234	2	Q8T5F2_9CNID
29	667	54.9	225	2	Q601Z5_9CNID
30	667	54.9	266	1	GPPL_GLASP
31	663	54.6	226	2	Q8T6U0_9CNID

32	660.5	54.4	227	2	Q66ND4_9CNID	Q66ND4_montastraea
33	655.5	54.0	226	2	Q66PW5_9CNID	Q66PW5_agaricia fr
34	655.5	54.0	227	2	Q95VT0_9CNID	Q95VT0_montastraea
35	654.5	53.9	234	2	Q8MU47_9CNID	Q8MU47_montastraea
36	652.5	53.7	225	2	Q7Z0W9_9CNID	Q7Z0W9_montastraea
37	649.5	53.5	227	2	Q5JUG7_9CNID	Q5JUG7_favia favus
38	649	53.5	225	2	Q601Z4_9CNID	Q601Z4_galaxea fas
39	642.5	52.9	225	2	Q6USK3_9CNID	Q6USK3_montastraea
40	639.5	52.7	227	2	Q5JUG8_9CNID	Q5JUG8_favia favus
41	639	52.6	224	2	Q8MU48_9CNID	Q8MU48_montastraea
42	631.5	52.0	225	2	Q8T5F1_9CNID	Q8T5F1_montastraea
43	620.5	51.1	230	2	Q66PW1_9CNID	Q66PW1_scolymia cu
44	610	50.2	223	2	Q6R8F5_9CNID	Q6R8F5_astrangia 1
45	608.5	50.1	231	1	RFP_PABAC	Q818F8_patasicyoni

## ALIGNMENTS

RESULT 1  
RFP DISSP ID RFP DISSP STANDARD; PRT; 225 AA.  
AC Q9U6Y8;  
DT 01-FEB-2005 (Rel. 46, Last sequence update)  
DT 01-FEB-2005 (Rel. 47, Last annotation update)  
DE Red fluorescent protein drp583 (Dereid).  
OS Discosoma sp. (See anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_Taxid=86600;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=99436614; PubMed=10504696; DOI=10.1038/13657;  
RA Matz M.V., Pradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G.,  
RA Markelov M.L., Lukyanov S.A.;  
RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";  
RL Nat. Biotechnol. 17:969-973(1999).  
RN (2)  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX PubMed=1101896; DOI=10.1038/81992;  
RA Wall M.A., Socolich M., Ranganathan R.;  
RT "The structural basis for red fluorescence in the tetrameric GFP homolog DsRed.";  
RL Nat. Struct. Biol. 7:1133-1138(2000).  
RN (3)  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND SUBUNIT.  
RX PubMed=11209050; DOI=10.1073/pnas.98.2.462;  
RA Yarbrough D., Wachter R.M., Kallio K., Matz M.V., Remington S.J.;  
RT Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-A resolution.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:462-467(2001).  
CC -I- FUNCTION: Thought to play a role in photoprotection of the coral's resident symbiotic microalgae's photosystems from photoinhibition caused by high light levels found near the surface of coral reefs. In deeper water, the fluorescence may be to convert blue light into longer wavelengths more suitable for use in photosynthesis by the microalgal symbionts.  
CC -I- BIOPHYSICOCHEMICAL PROPERTIES:  
CC Absorption: Abs(max)=558 nm;  
CC Note=Exhibits a smaller absorbance peak at 494 nm. The broad fluorescence emission spectrum peaks at 583 nm;  
CC -I- SUBUNIT: Homotrimer.  
CC -I- PTM: Contains a chromophore consisting of modified amino acid residues. The chromophore is formed by autocatalytic backbone condensation between Xaa-N and Gly-(N-2), oxidation of Tyr-(N-1) to dihydroxytyrosine, and formation of a double bond to the alpha-amino nitrogen of residue Xaa-N. Maturation of the chromophore requires nothing other than molecular oxygen.  
CC -I- BIOTECHNOLOGY: Fluorescent proteins have become a useful and ubiquitous tool for making chimeric proteins, where they function as a fluorescent protein tag. Typically they tolerate N- and C-

terminal fusion to a broad variety of proteins. They have been expressed in most known cell types and are used as a noninvasive fluorescent marker in living cells and organisms. They enable a wide range of applications where they have been functional as a cell lineage tracer, reporter of gene expression, or as a measure of protein-protein interactions.

-1- SIMILARITY: Belongs to the GFP family.

-----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-CC the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL, AF168419; AAF03369.1; -. mRNA.

CC PDB, 1G7K; X-ray; A/B/C/D=2-225.

CC PDB, 1G8X; X-ray; A/B/C/D=1-225.

CC InterPro; IPR011584; GFP related.

CC InterPro; IPR000786; Green\_fl\_protein.

CC Pfam; PF01353; GFP; 1.

CC PRINTS; PR01229; GFPUNRESIDENT.

CC ProDom; PD013756; Green\_fl\_protein; 1.

CC 3D-structure; Chromophore; Luminescence; Photoprotein.

CC MOD\_R1; 67 67 (2)-2,3-didehydrocytosine.

CC CROSBANK 66 68 2-immethyl-5-immidazolone (Gln-Gly).

CC SEQUENCE 225 AA; 25931 MW; FBR9A5369778F689 CRC64;

Query Match 100.0%; Score 1214; DB 1; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-96;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFVMEGTNGHGFEBEGEGRPYEGHNTVYLKVTGKGPLFPANDI 60  
 DB 1 MSSKNVKEFMRFVMEGTNGHGFEBEGEGRPYEGHNTVYLKVTGKGPLFPANDI 60  
 QY 61 LSPQFQYSKVVYVHPADIPYKKLSFPEGFKMERVWVNFEDGVVTVQDSSLQDGCFTY 120  
 DB 61 LSPQFQYSKVVYVHPADIPYKKLSFPEGFKMERVWVNFEDGVVTVQDSSLQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVMQKKTGMWEASTERYLPDGVLKGELHKALKLKDGGHYLVEFKSI 180  
 DB 121 KYKFIGVNPSPDGPVMQKKTGMWEASTERYLPDGVLKGELHKALKLKDGGHYLVEFKSI 180  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTGRRHHLFL 225  
 DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTGRRHHLFL 225

RESULT 2  
 Q66ND7\_9CNID PRELIMINARY; PRT; 236 AA.

QY Q66ND7\_9CNID PRELIMINARY; PRT; 236 AA.

DT 25-OCT-2004 (TREMBLrel. 28, Created)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
 DE Enhanced red fluorescent protein R+.

OS Discosoma sp. RC-2004.

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
 OC Discosomatidae; Discosoma.

OC NCBI\_TaxID=289055;

OX [1]

RP NUCLEOTIDE SEQUENCE.  
 RA Carter R.W., Gibbs P.D.L., Schmale M.C.;  
 RL "Cloning of Anthozoan Fluorescent Protein Genes";  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY679107; AAU0444.1; -. mRNA.  
 DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
 DR InterPro; IPR011584; GFP related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.  
 DR PRINTS; PR01229; GFPUNRESIDENT.  
 DR ProDom; PD013756; Green\_fl\_protein; 1.  
 SQ SEQUENCE 236 AA; 27032 MW; BB3844BE0682BEF0 CRC64;

Query Match 98.3%; Score 1193; DB 2; Length 236;  
 Best Local Similarity 98.2%; Pred. No. 3.9e-95;  
 Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFVMEGTNGHGFEBEGEGRPYEGHNTVYLKVTGKGPLFPANDI 60  
 DB 1 MSSKNVKEFMRFVMEGTNGHGFEBEGEGRPYEGHNTVYLKVTGKGPLFPANDI 60  
 QY 61 LSPQFQYSKVVYVHPADIPYKKLSFPEGFKMERVWVNFEDGVVTVQDSSLQDGCFTY 120  
 DB 61 LSPQFQYSKVVYVHPADIPYKKLSFPEGFKMERVWVNFEDGVVTVQDSSLQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVMQKKTGMWEASTERYLPDGVLKGELHKALKLKDGGHYLVEFKSI 180  
 DB 121 KYKFIGVNPSPDGPVMQKKTGMWEASTERYLPDGVLKGELHKALKLKDGGHYLVEFKSI 180  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTGRRHHLFL 225  
 DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTGRRHHLFL 225

RESULT 3  
 Q6KF85\_9CNID PRELIMINARY; PRT; 225 AA.

QY Q6KF85\_9CNID PRELIMINARY; PRT; 225 AA.

AC Q6KF85\_9CNID PRELIMINARY; PRT; 225 AA.

DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Orange fluorescent protein FP586.

OS Discosoma sp. JW-2002.

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
 OC Discosomatidae; Discosoma.

OC NCBI\_TaxID=208461;

OX [1]

RP NUCLEOTIDE SEQUENCE.  
 RA Wiedemann J., Girod A.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF545828; AA011987.1; -. mRNA.  
 DR HSRP; P42212; I89C.  
 DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
 DR InterPro; IPR011584; GFP related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.  
 DR PRINTS; PR01229; GFPUNRESIDENT.  
 DR ProDom; PD013756; Green\_fl\_protein; 1.  
 SQ SEQUENCE 225 AA; 25791 MW; E151D0E497AA23FA CRC64;

Query Match 98.3%; Score 1193; DB 2; Length 225;  
 Best Local Similarity 98.2%; Pred. No. 6.7e-95;  
 Matches 221; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFVMEGTNGHGFEBEGEGRPYEGHNTVYLKVTGKGPLFPANDI 60  
 DB 1 MSSKNVKEFMRFVMEGTNGHGFEBEGEGRPYEGHNTVYLKVTGKGPLFPANDI 60  
 QY 61 LSPQFQYSKVVYVHPADIPYKKLSFPEGFKMERVWVNFEDGVVTVQDSSLQDGCFTY 120  
 DB 61 LSPQFQYSKVVYVHPADIPYKKLSFPEGFKMERVWVNFEDGVVTVQDSSLQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVMQKKTGMWEASTERYLPDGVLKGELHKALKLKDGGHYLVEFKSI 180  
 DB 121 KYKFIGVNPSPDGPVMQKKTGMWEASTERYLPDGVLKGELHKALKLKDGGHYLVEFKSI 180  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTGRRHHLFL 225  
 DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTGRRHHLFL 225

RESULT 4  
 Q66ND8\_9CNID PRELIMINARY; PRT; 236 AA.

QY Q66ND8\_9CNID PRELIMINARY; PRT; 236 AA.

AC Q66ND8\_9CNID PRELIMINARY; PRT; 236 AA.

DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Red fluorescent protein R1.  
 OS Discosoma sp. RC-2004.  
 OC Eukaryota; Metazoa; Chordata; Anthozoa; Zoantharia; Corallimorpharia;  
 OC Discosomatidae; Discosoma.  
 NCBI\_TaxID=289055;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Carter R.W., Gibbs P.D.L., Schmale M.C.;  
 RT "Cloning of Anthozoan Fluorescent Protein Genes";  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DB EMBL; AY679106; AAU04443.1; mRNA.  
 DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
 DR InterPro; IPR011584; GFP related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.  
 DR PRINTS; PR01229; GFPUDRESCENT.  
 DR ProDom; PD013756; Green\_fl\_protein; 1.  
 SQ SEQUENCE 236 AA; 27042 MW; BB3844BCD6834EF3 CRC64;

Query Match 98.1%; Score 1191; DB 2; Length 236;  
 Best Local Similarity 97.8%; Pred. No. 1.1e-94;  
 Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRSKRVIVKEFMRFKVRMEGTGVNGHFEIEGEGGRPYEGHNTVTKLVTGKGPLPFAWDI 60  
 DB 1 MGSKRVIVKEFMRFKVRMEGTGVNGHFEIEGEGGRPYEGHNTVTKLVTGKGPLPFAWDI 60  
 QY 61 LSPFOYGSKVVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLDGCFIY 120  
 DB 61 LSPFOYGSKVVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLDGCFIY 120  
 QY 121 KYKFTGVNPPSDGPVWOKTGMGEASTERLYPRDGYLKGEIHKALKDKGCHYLVEFKSI 180  
 DB 121 KYKFTGVNPPSDGPVWOKTGMGEASTERLYPRDGYLKGEIHKALKDKGCHYLVEFKSI 180  
 QY 181 YNAKRPVOLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHFL 225  
 DB 181 YNAKRPVOLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHFL 225

# RESULT 5 Q5G7J7\_9CNID PRELIMINARY; PRT; 230 AA.

AC Q5G7J7\_9CNID PRELIMINARY; PRT; 230 AA.  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Red fluorescent protein.  
 GN Name-PF593;  
 OS Discosoma sp. SSAL-2000.  
 OC Eukaryota; Metazoa; Chordata; Anthozoa; Zoantharia; Corallimorpharia;  
 OC Discosomatidae; Discosoma.  
 NCBI\_TaxID=137428;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA MEDLINE=20434599; PubMed=10981720; DOI=10.1016/S0014-5793(00)01895-0;  
 RA Fradkov A.F., Chen Y., Ding L., Barsova E.V., Matz M.V.,  
 RA Lukyanov S.A.;  
 RT "Novel fluorescent protein from Discosoma coral and its mutants  
 RT possesses a unique far-red fluorescence.";  
 RL PNAS Lett. 479:127-130(2000).  
 DB EMBL; AF272711; AAC16224.1; mRNA.  
 DB HSP; Q9618; IGX.  
 DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
 DR InterPro; IPR011584; GFP related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.  
 DR PRINTS; PR01229; GFPUDRESCENT.  
 SQ SEQUENCE 230 AA; 26370 MW; 5215B1B436D67B51 CRC64;

Query Match 89.4%; Score 1085.5; DB 2; Length 230;  
 Best Local Similarity 88.1%; Pred. No. 1.3e-85;  
 Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;

QY 1 MRSKRVIVKEFMRFKVRMEGTGVNGHFEIEGEGGRPYEGHNTVTKLVTGKGPLPFAWDI 60  
 DB 1 MGSKRVIVKEFMRFKVRMEGTGVNGHFEIEGEGGRPYEGHNTVTKLVTGKGPLPFAWDI 60  
 QY 61 LSPFOYGSKVVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLDGCFIY 120  
 DB 61 LSPFOYGSKVVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLDGCFIY 120  
 QY 121 KYKFTGVNPPSDGPVWOKTGMGEASTERLYPRDGYLKGEIHKALKDKGCHYLVEFKSI 180  
 DB 121 KYKFTGVNPPSDGPVWOKTGMGEASTERLYPRDGYLKGEIHKALKDKGCHYLVEFKSI 180  
 QY 181 YNAKRPVOLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHFL 225  
 DB 181 YNAKRPVOLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHFL 225

# RESULT 6 Q5S3G8\_9CNID PRELIMINARY; PRT; 226 AA.

AC Q5S3G8\_9CNID PRELIMINARY; PRT; 226 AA.  
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
 DE Fluorescent protein raspberry.  
 OS Discosoma sp. LM-2004.  
 OC Eukaryota; Metazoa; Chordata; Anthozoa; Zoantharia; Corallimorpharia;  
 OC Discosomatidae; Discosoma.  
 NCBI\_TaxID=301246;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA PubMed=15556995; DOI=10.1073/pnas.0407752101;  
 RA Wang L., Jackson W.C., Steinbach P.A., Tsien R.Y.;  
 RT "Evolution of new nonantibody proteins via iterative somatic  
 RT hypermutation";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:16745-16749(2004).  
 DB EMBL; AY786536; AAV65486.1; mRNA.  
 DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
 DR InterPro; IPR011584; GFP related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.  
 DR PRINTS; PR01229; GFPUDRESCENT.  
 DR ProDom; PD013756; Green\_fl\_protein; 1.  
 SQ SEQUENCE 226 AA; 25513 MW; 159A8FBE2D453A5B CRC64;

Query Match 81.1%; Score 985; DB 2; Length 226;  
 Best Local Similarity 85.3%; Pred. No. 6.3e-77;  
 Matches 185; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

QY 5 KVIYKEMRFKVRMEGTGVNGHFEIEGEGGRPYEGHNTVTKLVTGKGPLPFAWDI 64  
 DB 6 EVIYKEMRFKVRMEGTGVNGHFEIEGEGGRPYEGHNTVTKLVTGKGPLPFAWDI 64  
 QY 65 FOYGSKVVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLDGCFIY 124  
 DB 66 CMYGSKVVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLDGCFIY 125  
 QY 125 IGYNPPSDGPVWOKTGMGEASTERLYPRDGYLKGEIHKALKDKGCHYLVEFKSI 184  
 DB 126 RGTNPPSDGPVWOKTGMGEASTERLYPRDGYLKGEIHKALKDKGCHYLVEFKSI 185  
 QY 185 KPVOLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRH 221  
 DB 186 KPVOLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRH 222

# RESULT 7 Q5S3G7\_9CNID PRELIMINARY; PRT; 226 AA.

AC Q5S3G7\_9CNID PRELIMINARY; PRT; 226 AA.

Query Match	80.1%	Score 972;	DB 2;	Length 226;
Best Local Similarly	83.4%	Pred. No. 8.4e-76;		
Matches 181; Conservative		8; Mismatches 28;	Indels 0;	Gaps 0.

RESULT 8	
Q66PV1 AC RTE	
ID Q66PV1 AC RTE PRELIMINARY;	PRT; 221 AA

DR PRODom; pD013/56; Green II protein; 1.  
SQ SEQUENCE 221 AA; 25103 MW; 3BEA0B848C8F5DA4 CRC64;

```

6 NIKEFNFKFKVMEGTANGHEFELEBEGGEGRPYEGHNTVKLVTKGGLPLPFANDILSPQF 65
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
2 SVIAKOMTYKVMSGTANGHYFFEBGDGCKPKYEEBQTVLKTVTKGSLPLPFANDILSPQF 61

```

186 PVQLPGYVYDSDKLITSHNEDYTIIVEQYRTEGRHLL 223  
||::|::|::|::|::|::|::|::|::|::|::|::|::|  
182 PVRMPGYHYDRKLDVTNNHKDYTSVEQCCELSIARPEL 219

Query Match	65.7%	Score 798	DB 2	Length 221
Best Local Similarity	69.0%	Pred. No. 8.9e-61		
Matches 145; Conservative	24	Mismatches 41	Indels 0	Gaps 0

RESULT 10	
066PV0_ACRMI	
066PV0_ACRMI	PRELIMINARY;
066PV0_	PRT;
25-OCT-2004 (TTEMBLrel. 28, Created)	221 AA

CC 25-OCY-2004 (Tremblrel. 28, Last sequence update)  
DT 25-OCY-2004 (Tremblrel. 28, Last annotation update)  
DE Chromoprotein.  
OS Acropora millepora (Coral).  
OC Eukaryota; Metazoa; Anthozoa; Zoantharia; Scleractinia;  
OC Scleractinia; Acroporidae; Acropora.  
OX NCBI\_TaxID=45364;  
RN [1]  
RA NCULEOTIDE SEQUENCE.  
RP Allvea N.O., Melashkevitch E.A., Field S.F., Matz M.V.;  
RT "Survey of coral GFP-like proteins."  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL, AY645075, LAU06854.1; -. mRNA.  
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
DR InterPro; IPR011584; GFP-related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP. 1.  
DR PRINTS; PR01229; GFPUDORESCENT.  
DR PRODOM; PD013756; Green\_fl\_protein. 1.  
DR SEQUENCE 221 AA; 24967 MW; 546D2CFPAF2552F2 CRC64;  
SQ  
Query Match 65.7%; Score 797; DB 2; Length 221;  
Best Local Similarity 68.6%; Pred. No. 1.le-60;  
Matches 144; Conservative 26; Mismatches 40; Indels 0; Gaps 0  
DB 2 SVIAKQMTYKVMSTGVNGHYFEVGDGKKEYPEGQTKLITVTKGGLPFAMDILSPQC 61  
QY 6 NVIKFPMFKYRMESTGVNGHEFEIIGEGEGRPYEGHNTYKLYTKGGLPFAMDILSPQ 65  
DB 2 SVIAKQMTYKVMSTGVNGHYFEVGDGKKEYPEGQTKLITVTKGGLPFAMDILSPQC 61  
QY 66 QYGSKVTYKHPADIDYKLSPEEGFKMRVNNFEDGGVTVTVTQDSDGGCFYKVKFI 125  
DB 62 QYGSIPFTKYPEPIDDYKQSPFEGTWRIRNPFEGACVTSNDSISGNCYTHVKS 121  
QY 126 GVNFPSPDPVMOCKTMGWEASTERYLPBDGVKGEIHKALKIDGGHYLVFKFSIYMAK 185  
DB 122 GLNFPNPGPVMOCKTQGWEPNTERLPARGCMILGNPPALNKLGGSHYLCFPTTYKAKK 181  
QY 186 PVQLPGYYVDSKLDITSHEDYTYVEQYE 215  
DB 182 PVKMGYHYVDRKLDVTNHNKDYTSVEQCE 211  
RESULT 11  
NFCP GONTE  
ID \_NFCP GONTE STANDARD; PRT; 221 AA.  
AC 09SP04;  
DT 01-FEB-2005 (Rel. 46, Created)  
DT 01-FEB-2005 (Rel. 46, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE GFP-like non-fluorescent chromoprotein (GfCP).  
OS Acropora tenuis (Coral).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
OC Fungiina; Poritidae; Goniopora.  
OX NCBI\_TaxID=75301;  
RN [1]  
RA NCULEOTIDE SEQUENCE, SUBUNIT, AND MUTAGENESIS OF ASN-142.  
RP MEDLINE=21538666; PubMed=11682051; DOI=10.1016/S0014-5793(01)02930-1;  
RX Gurekaya N.G., Fridkov A.F., Terskikh A., Matz M.V., Labas Y.A.,  
RA Matrynov V.I., Yanushkevich Y.G., Lukyanov K.A., Lukyanov S.A.;  
RT "GFP-like chromoproteins as a source of far-red fluorescent  
RT proteins."  
RL FEBS Lett. 507:16-20(2001).  
CC -1- FUNCTION: Non-fluorescent pigment protein that is lilac in color.  
CC -1- BIOCHEMICAL PROPERTIES:  
CC Absorption:  
CC Abs(max)=580 nm;  
CC -1- SUBUNIT: Homotrimer.  
CC -1- PTM: Contains a chromophore consisting of modified amino acid  
CC residues. The chromophore is formed by autocatalytic backbone  
CC condensation between Xaa-N and Gly-(N+2), oxidation of Tyr-(N+1)  
CC to dihydroxyrosine, and formation of a double bond to the alpha-  
CC amino nitrogen of residue Xaa-N. Maturation of the chromophore  
CC requires nothing other than molecular oxygen. The precise

CC BIOCHEMISTRY OF THE TYROSINE HAS NOT BEEN DETERMINED.

CC BIOTECNOLOGY: Fluorescent proteins have become a useful and

CC ubiquitous tool for making chimeric proteins, where they function

CC as a fluorescent protein tag. Typically they tolerate N- and C-

CC terminal fusion to a broad variety of proteins. They have been

CC expressed in most known cell types and are used as a noninvasive

CC fluorescent marker in living cells and organisms. They enable a

CC wide range of applications where they have functioned as a cell

CC lineage tracer, reporter of gene expression, or as a measure of

CC protein-protein interactions.

CC -1- MISCELLANEOUS: Fluorescence excitation of the Ser-142 mutant is at

CC 589 nm and emission at 615 nm.

CC -1- SIMILARITY: Belongs to the GFP family.

-----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

-----

CC EMBL; AF383156; AL27542.1; -, mRNA.

DR HSSP; Q906Y8; IGGX.

DR GO; GO:0018995; C:host; ISS.

DR GO; GO:0010117; P:photoprotection; ISS.

DR InterPro; IPR00786; GFP\_related.

DR InterPro; IPR01584; GFP\_protein.

DR Pfam; PF01353; GFP\_1.

DR PRINTS; PR01229; GFP\_LUORESCENT.

KW Chromophore; luminescence; Photoprotein.

FT MOD\_RES 63 63 2,3-dihydrotyrosine (By similarity).

FT CROSSLINK 62 64 2-iminomethyl-5-imidazolinone (Gln-Gly)

FT (By similarity).

FT MUTAGEN 142 142 N-5: Produces a fluorescent form.

SQ SEQUENCE 221 AA; 24918 MW; 93F9PB5C2003CB4 CRC64;

Query Match 65.3%; Score 793; DB 1; Length 221;

Best Local Similarity 68.6%; Pred. No. 2,4e-60;

Matches 144; Conservative 25; Mismatches 41; Indels 0; Gaps 0,

QY 6 NVIKKFMKFKVMEGTGVNGHEFEIIBGGEGRPYEGHNTVKLKVTKGGLPFAMIDLSQF 65  
:::  
DB 2 SVIAKQMYYKYVMGSTVNGHYFEVGDDGKKRPGEGEQTKLVTKIGGPLPFAMIDLSPOS 61  
:::  
QY 66 QYGSKNVYKHAADIPDYKLKSPFGSGFKMERIMNPDGCVVMTDSSLSDCGFYTKVEFI 125  
||| :  
DB 62 QYGSIPPKYKEPIDIPDYKOSFPBRYTERTIMNPBDGANVCIVSNDSISIQNCFIYNKFS 121  
||| :  
QY 126 GVNPSPDSPVNMOKTMGEASTERLYPRPDGYLKGIEHKALKDKGHVLVEFKSIYMAKK 185  
:||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:  
DB 122 GLNPPNPSPVNMOKTIQGWEPTERTLPARDGMIGNFMALKLBSGGHILCEFKSYTKAKK 181  
:||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:  
QY 186 PVQLPGYVVVNSKLDITSHNEDYIVEOYE 215  
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:  
DB 182 FVQMPGHHYVDKLDVTNHNIDYSVEGCE 211  
||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 12

O66PU9 9CNID PRELIMITARY; PTR; 221 AA.

ID O66PU9 9CNID PRELIMITARY; PTR; 221 AA.

AC O66PU9 9CNID PRELIMITARY; PTR; 221 AA.

DT 25-OCT-2004 (TREMBLrel. 28, Created)

DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)

D8 Chromoprotein.

OS Acropora hyacinthus.

OC Eukaryota; Metazoa; Chordata; Anthozoa; Zoantharia; Scleractinia;

OC Astrocenina; Actinoptera; Actinoptera.

OX NCBI\_TaxID=55974;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Aliava N.O., Melashkevitch B.A., Field S.F., Matz M.V.;

RT "Survey of coral GFP-like proteins."

Submitted (JUN-2004) to the EMBL/Genbank/DDBJ databases.



DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.

DR PRINTS; PRO1229; GFP.1.  
 DR PRODOM; PD013756; Green\_fl\_protein; 1.

DR SEQUENCE 227 AA; 25640 MW; 568639306F44B244 CRC64;

Query Match 61.4%; Score 745; DB 2; Length 227;

Best Local Similarity 60.8%; Pred. No. 3.5e-56;  
 Matches 135; Conservative 33; Mismatches 54; Indels 0; Gaps 0;

QY 1 MSSKAVIKKPKRMKMGTVNGHFELEGEGRPVYGHNTVTKLTKGGLPAPWDI 60

DB 1 MSWSKSVIKKEMLDLHLEGTNGHFELEGKKGKGNPGNTVLTLETKGGLPFGMH 60

QY 61 LSPQFGSKVYKHPADIPDYKLSFPPGFKERVMNPFEDGCVTVTQDSSLQDCFTY 120

DB 61 LCPQFGYGNKAFVHHDPIDPDKLSFPPGYTWERSMHPEDGGLCITINDISLTGNCFTY 120

QY 121 KYKFTGVNPPSDGPNVQKKTGMWEASTERLYPRDGLKGEIHKALKKDGHTLVFESKI 180

DB 121 DKFTGLNPPNGPVVQKKTGMWEPSTERLYPRDGLIGDIHHLTVAGGHTVCDIKTV 180

QY 181 YMAKPVQLPFGYVDSKLDITSHNEDYTVQYERTGRHH 222

DB 181 YRAKPKVAKMFGYHYVDPKLVIRSDKPFMKVEHEHIAVARHH 222

RESULT 15  
 GFFL\_DIST STANDARD; PRT; 232 AA.

AC Q9U6Y7;  
 DT 01-FEB-2005 (Rel. 46, Created)

DT 01-FEB-2005 (Rel. 46, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DB GFP-like fluorescent chromoprotein depp483.

DE Discosoma striata (Striped mushroom).

OC Bkaryocia; Metazoa; Chordata; Anthozoa; Zoantharia; Corallimorpharia;  
 OC Discosomatidae; Discosoma.

NCBI\_TaxId=105400;

NUCLEOTIDE SEQUENCE, FUNCTION, AND TISSUE SPECIFICITY.

RA MEDLINE=99456514; PubMed=10504696; DOI=10.1038/136571;

RA Matz M.V., Prackov A.P., Labas Y.A., Savitsky A.P., Zaraisky A.G.,  
 Markelov M.L., Lukyanov S.A.;

RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";  
 Nat. Biotechnol. 17:969-973(1999).

CC -1- FUNCTION: Pigment protein that is green in color.

CC -1- BIOPHYSICO-CHEMICAL PROPERTIES:  
 Absorption:  
 Abs(max)=443 nm;  
 Note=Has a strong fluorescence emission spectrum which peaks at  
 483 nm;

CC -1- TISSUE SPECIFICITY: Oral disk.

CC -1- PTM: Contains a chromophore consisting of modified amino acid  
 residues. The chromophore is formed by autocatalytic backbone  
 condensation between Xaa-N and Gly-(N+2), oxidation of Tyr-(N+1)  
 to dihydroxytyrosine, and formation of a double bond to the alpha-  
 amino nitrogen of residue Xaa-N. Maturation of the chromophore  
 requires nothing other than molecular oxygen. The precise  
 stereochemistry of the tyrosine has not been determined.

CC -1- BIOTECHNOLOGY: Fluorescent proteins have become a useful and  
 ubiquitous tool for making chimeric proteins, where they function  
 as a fluorescent protein tag. Typically they tolerate N- and C-  
 terminal fusion to a broad variety of proteins. They have been  
 expressed in most known cell types and are used as a noninvasive  
 fluorescent marker in living cells and organisms. They enable a  
 wide range of applications where they have functioned as a cell  
 lineage tracer, reporter of gene expression, or as a measure of  
 protein-protein interactions.

CC -1- SIMILARITY: Belongs to the GFP family.

CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.

CC -----  
 CC EMBL; AF168420; AAF03370.1; -; mRNA.

CC HSSP; Q9U6Y8; 1GGX.

CC GO; GO:0008218; P:bioluminescence; IS8.

CC InterPro; IPR011584; GFP-related.

CC InterPro; IPR000786; Green\_fl\_protein.

CC Pfam; PF01353; GFP; 1.

CC PRINTS; PRO1229; GFP.1.

CC PRODOM; PD013756; Green\_fl\_protein; 1.

CC Chromophore; Luminescence; Photoprotein.

CC MOD\_RES 67 67 2,3-dihydroxytyrosine (By similarity).  
 FT 2-iminomethyl-5-imidazolone (Gln-Gly)  
 FT CROSSLINK 66 68 (By similarity).

CC SEQUENCE 232 AA; 26435 MW; AAE18BE8283CE4D CRC64;

Query Match 60.1%; Score 729.5; DB 1; Length 232;

Best Local Similarity 59.1%; Pred. No. 7.8e-55;  
 Matches 133; Conservative 38; Mismatches 53; Indels 1; Gaps 1;

QY 1 MSSKAVIKKPKRMKMGTVNGHFELEGEGRPVYGHNTVTKLTKGGLPAPWDI 60

DB 1 MSCKSVIKKEMLDLHLEGTNGHFELEGKKGKGNPGNTVLTLETKGGLPFGMH 60

QY 61 LSPQFGSKVYKHPADIPDYKLSFPPGFKERVMNPFEDGCVTVTQDSSLQDCFTY 120

DB 61 LCPQFGYGNKAFVHHDPIDPDKLSFPPGYTWERSMHPEDGGLCITINDISLTGNCFTY 120

QY 121 KYKFTGVNPPSDGPNVQKKTGMWEASTERLYPRDGLKGEIHKALKKDGHTLVFESKI 180

DB 121 DKFTGLNPPNGPVVQKKTGMWEPSTERLYPRDGLIGDIHHLTVAGGHTVCDIKTV 180

QY 181 YMAK-K-PVQLPFGYVDSKLDITSHNEDYTVQYERTGRHH 224

DB 181 YRAKKAALKMFGYHYVDTLKVIRSDKPFMKVEHEHIAVARHH 225

Search completed: January 11, 2006, 02:04:22  
 Job time : 74 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 11, 2006, 01:59:57 (Search time 23 Seconds  
(without alignments)  
808.784 Million cell updates/sec

Title: US-10-006-922a-12

Perfect score: 1214

Sequence: 1 MRSSKNVYKEFMRFKVRMEG.....EDTYVEQYERTGRHHLFL 225

Scoring table: BLOSUM62

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/6\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/H\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/RE\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	US-09-866-538-12	Sequence 12, Appl
2	1214	100.0	225	US-09-865-281-12	Sequence 12, Appl
3	1210	99.7	240	US-10-152-286-2	Sequence 2, Appl
4	681	56.1	225	US-10-244-779-2	Sequence 2, Appl
5	485.5	40.0	238	US-09-277-716-16	Sequence 16, Appl
6	485.5	40.0	238	US-09-609-161B-16	Sequence 16, Appl
7	485.5	40.0	238	US-09-626-581D-65	Sequence 65, Appl
8	485.5	40.0	238	US-09-415-765B-65	Sequence 65, Appl
9	485.5	40.0	238	US-09-626-580C-65	Sequence 65, Appl
10	473.5	39.0	224	US-09-977-897-3	Sequence 3, Appl
11	473.5	39.0	224	US-09-977-897-13	Sequence 13, Appl
12	473.5	39.0	225	US-09-977-897-14	Sequence 14, Appl
13	473.5	39.0	226	US-09-977-897-15	Sequence 15, Appl
14	473.5	39.0	227	US-09-977-897-16	Sequence 16, Appl
15	473.5	39.0	228	US-09-977-897-17	Sequence 17, Appl
16	473.5	39.0	229	US-09-977-897-18	Sequence 18, Appl
17	473.5	39.0	230	US-09-977-897-19	Sequence 19, Appl
18	473.5	39.0	231	US-09-977-897-20	Sequence 20, Appl
19	473.5	39.0	232	US-09-977-897-21	Sequence 21, Appl
20	473.5	39.0	233	US-09-977-897-22	Sequence 22, Appl
21	473.5	39.0	234	US-09-977-897-9	Sequence 9, Appl
22	473.5	39.0	234	US-09-977-897-23	Sequence 23, Appl
23	473.5	39.0	235	US-09-977-897-8	Sequence 8, Appl
24	473.5	39.0	235	US-09-977-897-24	Sequence 24, Appl
25	473.5	39.0	236	US-09-977-897-7	Sequence 7, Appl
26	473.5	39.0	236	US-09-977-897-25	Sequence 25, Appl
27	473.5	39.0	237	US-09-977-897-6	Sequence 6, Appl

28	473.5	39.0	237	US-09-977-897-26	Sequence 26, Appl
29	473.5	39.0	238	US-09-277-716-32	Sequence 32, Appl
30	473.5	39.0	238	US-09-609-161B-32	Sequence 32, Appl
31	473.5	39.0	238	US-09-839-650-3	Sequence 3, Appl
32	473.5	39.0	238	US-09-977-897-5	Sequence 5, Appl
33	473.5	39.0	238	US-09-977-897-27	Sequence 27, Appl
34	473.5	39.0	238	US-10-652-703A-3	Sequence 3, Appl
35	473.5	39.0	239	US-09-977-897-2	Sequence 2, Appl
36	473	39.0	219	US-09-977-897-4	Sequence 4, Appl
37	473	39.0	231	US-09-977-897-12	Sequence 12, Appl
38	473	39.0	232	US-09-977-897-11	Sequence 11, Appl
39	473	39.0	233	US-09-977-897-10	Sequence 10, Appl
40	465.5	38.3	239	US-10-021-818A-4	Sequence 4, Appl
41	464.5	38.3	238	US-10-021-818A-2	Sequence 2, Appl
42	215.5	17.8	238	US-09-023-946B-23	Sequence 23, Appl
43	214.5	17.7	238	US-09-023-946B-28	Sequence 28, Appl
44	213.5	17.6	240	US-09-129-192C-49	Sequence 49, Appl
45	212.5	17.5	238	US-08-818-604-32	Sequence 32, Appl

## ALIGNMENTS

RESULT 1  
US-09-866-538-12  
Sequence 12, Application US/09866538  
Patent No. 6852849  
GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSIEN, Roger  
APPLICANT: Campbell, Robert  
TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
FILE REFERENCE: REGENT1530-2  
CURRENT APPLICATION NUMBER: US/09/866, 538  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 12  
LENGTH: 225  
TYPE: PRT  
ORGANISM: *Dicobosoma* sp.  
US-09-866-538-12

Query Match 100.0%; Score 1214; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 3e-132;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRSSKNVYKEFMRFKVRMEGTGNGHEPEIEGEGRPRYEGHNTVYKLTGPGPLPFAMD	60
DB	1	MRSSKNVYKEFMRFKVRMEGTGNGHEPEIEGEGRPRYEGHNTVYKLTGPGPLPFAMD	60
QY	61	LSPOFGYGSRYVYVHPADIPDYKLSFPEGFKWERNVNFEDGGVTVYVQSSIQDGCFTY	120
DB	61	LSPOFGYGSRYVYVHPADIPDYKLSFPEGFKWERNVNFEDGGVTVYVQSSIQDGCFTY	120
QY	121	KVFETGVPFSDGPPVWQKTMGWEASIERLYPRDGVYKSGIIRALKKDGSHYVERKSI	180
DB	121	KVFETGVPFSDGPPVWQKTMGWEASIERLYPRDGVYKSGIIRALKKDGSHYVERKSI	180
QY	181	YMAKPPVQLPGYVYVDSKLDITSNEDYTIIVEQYERTGRHHLFL	225
DB	181	YMAKPPVQLPGYVYVDSKLDITSNEDYTIIVEQYERTGRHHLFL	225

RESULT 2  
US-09-865-291-12  
Sequence 12, Application US/09865291  
Patent No. 6900304  
GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSIEN, Roger  
APPLICANT: TING, Alice  
APPLICANT: ZHANG, Jin

```

; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discozona sp.
US-09-865-291-12

Query Match      100.0%; Score 1214; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 3e-132;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRFVYRMEGTNGHGFELIEGEGRPYEGHNTVTKLKTGGPLPFAMDI 60
DB 1 MRSSKNVKEFRFVYRMEGTNGHGFELIEGEGRPYEGHNTVTKLKTGGPLPFAMDI 60
QY 61 LSPQOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFDGCVVTVQDSSLQDGCFTY 120
DB 61 LSPQOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFDGCVVTVQDSSLQDGCFTY 120
QY 121 KYKFIGVNPSPDGPVMOQKTMGWEASTERLYPRDGVKGEIHKALKDKDGHYLVFPKSI 180
DB 121 KYKFIGVNPSPDGPVMOQKTMGWEASTERLYPRDGVKGEIHKALKDKDGHYLVFPKSI 180
QY 121 KYKFIGVNPSPDGPVMOQKTMGWEASTERLYPRDGVKGEIHKALKDKDGHYLVFPKSI 180
DB 121 KYKFIGVNPSPDGPVMOQKTMGWEASTERLYPRDGVKGEIHKALKDKDGHYLVFPKSI 180
QY 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIEQYERTGRHHLFL 225
DB 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIEQYERTGRHHLFL 225

RESULT 3
US-10-152-296-2
; Sequence 2, Application US/10152296
; Patent No. 6723537
; GENERAL INFORMATION:
; APPLICANT: Peelle, Beau
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells
; FILE REFERENCE: 021044-000110US
; CURRENT APPLICATION NUMBER: US/10/152,296
; CURRENT FILING DATE: 2002-12-10
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,871
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mammalian
; OTHER INFORMATION: codon-optimized variant (BRED) of Discozona sp.
; OTHER INFORMATION: "red" red fluorescent protein (RFP)
US-10-152-296-2

Query Match      99.7%; Score 1210; DB 2; Length 240;
Best Local Similarity 99.6%; Pred. No. 9.5e-132;
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRFVYRMEGTNGHGFELIEGEGRPYEGHNTVTKLKTGGPLPFAMDI 60
DB 2 VRSSKNVKEFRFVYRMEGTNGHGFELIEGEGRPYEGHNTVTKLKTGGPLPFAMDI 61
QY 61 LSPQOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFDGCVVTVQDSSLQDGCFTY 120
DB 62 LSPQOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFDGCVVTVQDSSLQDGCFTY 121
QY 121 KYKFIGVNPSPDGPVMOQKTMGWEASTERLYPRDGVKGEIHKALKDKDGHYLVFPKSI 180
DB 122 KYKFIGVNPSPDGPVMOQKTMGWEASTERLYPRDGVKGEIHKALKDKDGHYLVFPKSI 181
```

```

QY 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIEQYERTGRHHLFL 225
DB 182 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIEQYERTGRHHLFL 226

RESULT 4
US-10-244-779-2
; Sequence 2, Application US/10244779
; Patent No. 6933375
; GENERAL INFORMATION:
; APPLICANT: Falkowski, Paul
; APPLICANT: Sun, Yi
; APPLICANT: Gorbunov, Maxim
; APPLICANT: Wyman, Kevin
; APPLICANT: Chen, Yi-Bu
; TITLE OF INVENTION: mRFP Encoding Nucleic Acids,
; TITLE OF INVENTION: Polypeptides, Antibodies and Methods of Use Thereof
; FILE REFERENCE: Rut 00-002305
; CURRENT APPLICATION NUMBER: US/10/244,779
; CURRENT FILING DATE: 2002-12-19
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Monastrea cavernosa
US-10-244-779-2

Query Match      56.1%; Score 681; DB 2; Length 225;
Best Local Similarity 57.2%; Pred. No. 1.7e-70;
Matches 123; Conservative 36; Mismatches 56; Indels 0; Gaps 0;

QY 6 NVKEFMKFKVMEGTNGHGFELIEGEGRPYEGHNTVTKLKTGGPLPFAMDI 65
DB 2 SVKSVMKIKLRMDGI VNGHKFMITGEGKPEGTHTIILKVEGGLPFAYDILTAA 61
QY 66 QYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFDGCVVTVQDSSLQDGCFTY 125
DB 62 QYGNRVFTKYPKDI PDYKOSFPEGYSWERSMTEDGCVTVTSIDKLEGGCFYERIPY 121
QY 126 GVNFPSPDGPVMOQKTMGWEASTERLYPRDGVKGEIHKALKDKDGHYLVFPKSI 185
DB 122 GVNFPSPDGPVMOQKTMGWEASTERLYPRDGVKGEIHKALKDKDGHYLVFPKSI 181
QY 186 PVQLPGYYVDSKLDITSHNEDYTVIEQYERTGR 220
DB 182 GVVLPEYHFDHRIEILSHDKDYTVVEYENAVAR 216

RESULT 5
US-09-277-716-16
; Sequence 16, Application US/09277716A
; Patent No. 6232107
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
; CURRENT APPLICATION NUMBER: US/09/277,716A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 238
```

TYPE: PRT  
ORGANISM: Renilla mulleri  
FEATURE:  
OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)  
US-09-277-716-16

Query Match 40.0%; Score 485.5; DB 2; Length 238;  
Best Local Similarity 44.5%; Pred. No. 8.5e-48;  
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;

QY 4 SKNVK-----EFMRKVMEGTVNGHEFEIEGEGRGREYEGHNTYKLVKTKGGPLPFAM 58  
DB 2 SKQILKNTCLQGVMSKYKVLBEIYNNHVFYTMGCGGNILFGNQLVQIVTKGAPLPFAF 61  
QY 59 DILSPQFGSKYVKGADIPDYKLSPEEGFKMERVNFEDGVTVTODSSLQDGGCF 118  
DB 62 DIVSPAFQYGNRTFTKYPNDISDYFIQSFPAGFMERTLRYEDGGIVLRSDINLIDKRF 121  
QY 119 IYKVKRIGVNFPSDGVPMQKTMGWEASTERYLPDGVKGEIHKALKDKGGHYLVEPK 178  
DB 122 VYRVEKGSNFPDGDGVPMQKTLIGIEPSFEAMMNGVLVEVILVYKLSNGKYYSCMVK 181  
QY 179 SIYMAKPV-QLPGYVYVDSKLDITSHNEDYTVVEGYE 215  
DB 182 TLMKSGVYKPEPSYHFIOHRLKLT-YVEDGGFVEQHE 218

RESULT 6  
US-09-609-161B-16  
Sequence 16, Application US/09609161B  
Patent No. 6436682  
GENERAL INFORMATION:  
APPLICANT: Bryan, Bruce  
APPLICANT: Stent-Gyorgyi, Christopher  
APPLICANT: PROLUME, LTD.  
TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC  
TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG  
FILE REFERENCE: 24729-121B  
CURRENT APPLICATION NUMBER: US/09/609,161B  
CURRENT FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: 09/277,716  
PRIOR FILING DATE: 1999-03-26  
PRIOR APPLICATION NUMBER: 60/102,939  
PRIOR FILING DATE: 1998-10-01  
PRIOR APPLICATION NUMBER: 60/089,367  
PRIOR FILING DATE: 1998-06-15  
PRIOR APPLICATION NUMBER: 60/079,624  
PRIOR FILING DATE: 1998-03-27  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 16  
LENGTH: 238  
TYPE: PRT  
ORGANISM: Renilla mulleri  
FEATURE:  
OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)  
US-09-609-161B-16

Query Match 40.0%; Score 485.5; DB 2; Length 238;  
Best Local Similarity 44.5%; Pred. No. 8.5e-48;  
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;  
QY 4 SKNVK-----EFMRKVMEGTVNGHEFEIEGEGRGREYEGHNTYKLVKTKGGPLPFAM 58  
DB 2 SKQILKNTCLQGVMSKYKVLBEIYNNHVFYTMGCGGNILFGNQLVQIVTKGAPLPFAF 61  
QY 59 DILSPQFGSKYVKGADIPDYKLSPEEGFKMERVNFEDGVTVTODSSLQDGGCF 118  
DB 62 DIVSPAFQYGNRTFTKYPNDISDYFIQSFPAGFMERTLRYEDGGIVLRSDINLIDKRF 121  
QY 119 IYKVKRIGVNFPSDGVPMQKTMGWEASTERYLPDGVKGEIHKALKDKGGHYLVEPK 178

DB 122 VYRVEKGSNFPDGDGVPMQKTLIGIEPSFEAMMNGVLVEVILVYKLSNGKYYSCMVK 181  
QY 179 SIYMAKPV-QLPGYVYVDSKLDITSHNEDYTVVEGYE 215  
DB 182 TLMKSGVYKPEPSYHFIOHRLKLT-YVEDGGFVEQHE 218

RESULT 7  
US-09-626-581D-65  
Sequence 65, Application US/09626581D  
Patent No. 6548249  
GENERAL INFORMATION:  
APPLICANT: Anderson, David  
TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide  
FILE REFERENCE: A-66900-3/RMS  
CURRENT APPLICATION NUMBER: US/09/626,581D  
CURRENT FILING DATE: 2000-07-27  
PRIOR APPLICATION NUMBER: 09/169,015  
PRIOR FILING DATE: 1998-10-08  
PRIOR APPLICATION NUMBER: 09/415,765  
PRIOR FILING DATE: 1999-10-08  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 65  
LENGTH: 238  
TYPE: PRT  
ORGANISM: Renilla mulleri  
US-09-626-581D-65

Query Match 40.0%; Score 485.5; DB 2; Length 238;  
Best Local Similarity 44.5%; Pred. No. 8.5e-48;  
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;  
QY 4 SKNVK-----EFMRKVMEGTVNGHEFEIEGEGRGREYEGHNTYKLVKTKGGPLPFAM 58  
DB 2 SKQILKNTCLQGVMSKYKVLBEIYNNHVFYTMGCGGNILFGNQLVQIVTKGAPLPFAF 61  
QY 59 DILSPQFGSKYVKGADIPDYKLSPEEGFKMERVNFEDGVTVTODSSLQDGGCF 118  
DB 62 DIVSPAFQYGNRTFTKYPNDISDYFIQSFPAGFMERTLRYEDGGIVLRSDINLIDKRF 121  
QY 119 IYKVKRIGVNFPSDGVPMQKTMGWEASTERYLPDGVKGEIHKALKDKGGHYLVEPK 178  
DB 122 VYRVEKGSNFPDGDGVPMQKTLIGIEPSFEAMMNGVLVEVILVYKLSNGKYYSCMVK 181  
QY 179 SIYMAKPV-QLPGYVYVDSKLDITSHNEDYTVVEGYE 215  
DB 182 TLMKSGVYKPEPSYHFIOHRLKLT-YVEDGGFVEQHE 218

RESULT 8  
US-09-415-765B-65  
Sequence 65, Application US/09415765B  
Patent No. 6548632  
GENERAL INFORMATION:  
APPLICANT: Anderson, David  
TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide  
FILE REFERENCE: A66900-1/RMS/AMS  
CURRENT APPLICATION NUMBER: US/09/415,765B  
CURRENT FILING DATE: 1999-10-08  
PRIOR APPLICATION NUMBER: 09/169,015  
PRIOR FILING DATE: 1998-10-08  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 65  
LENGTH: 238  
TYPE: PRT  
ORGANISM: Renilla mulleri  
US-09-415-765B-65

Query Match 40.0%; Score 485.5; DB 2; Length 238;  
QY 4 SKNVK-----EFMRKVMEGTVNGHEFEIEGEGRGREYEGHNTYKLVKTKGGPLPFAM 58  
DB 2 SKQILKNTCLQGVMSKYKVLBEIYNNHVFYTMGCGGNILFGNQLVQIVTKGAPLPFAF 61  
QY 59 DILSPQFGSKYVKGADIPDYKLSPEEGFKMERVNFEDGVTVTODSSLQDGGCF 118  
DB 62 DIVSPAFQYGNRTFTKYPNDISDYFIQSFPAGFMERTLRYEDGGIVLRSDINLIDKRF 121  
QY 119 IYKVKRIGVNFPSDGVPMQKTMGWEASTERYLPDGVKGEIHKALKDKGGHYLVEPK 178  
DB 122 VYRVEKGSNFPDGDGVPMQKTLIGIEPSFEAMMNGVLVEVILVYKLSNGKYYSCMVK 181  
QY 179 SIYMAKPV-QLPGYVYVDSKLDITSHNEDYTVVEGYE 215  
DB 182 TLMKSGVYKPEPSYHFIOHRLKLT-YVEDGGFVEQHE 218





	Matches	97, Conservative	40, Mismatches	73, Indels	3, Gaps	3, Indels
Qy	5	KNV-IKEMRKVKMEGVNNGHEPIEESGSGRPVEGHNITYKALKTKCGRPADINDSP	63			
Db	8	KNTELKETMSAKAVEGIVNNHNVSMSEFGKGNVLFGQLMQIRYTKGGLPFADIVSI	67			
Qy	64	QFOYGSKYVVGHPADIPDYKKLSFPEEGEKMERVNMVFEDGGVTVTQDGLDGCETIRAK	123			
Db	68	AFQYGNRTFTKPPDDIADYFVQSGFPAGFYFYERNLRFEDGALVDIRSDLSDDKHYKVE	127			
Qy	124	PIGVNPFSDGQVWQKTMGMENASTERLYPRDQVTKGRIHAKLKLKDGCHYLVERKSIYMA	183			
Db	128	YRGGFPSSNGVPMOKALIGMEPSPFVVVYVMSGVLVGVDLVYKLESNGYVSCNMKTFYRS	187			
Qy	184	KKPV-QLPGYVYDSDKDITSHNEPDTIVEQYE	215			
Db	188	KGVYKEPEPEYHFIHRLKLT-YVEBGSVVEQHE	219			

Search completed: January 11, 2006, 02:05:23  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 02:00:37 ; Search time 62 Seconds  
(without alignment)  
1516.316 Million cell updates/sec

Title: US-10-006-922a-12

Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVMEG.....EDTYVEYERTGHHFL 225

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA\_Main:\*

- 1: /cgn2\_6/prodata/1/pubppa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/prodata/1/pubppa/US08\_PUBCOMB.pep:\*
- 3: /cgn2\_6/prodata/1/pubppa/US09\_PUBCOMB.pep:\*
- 4: /cgn2\_6/prodata/1/pubppa/US10A\_PUBCOMB.pep:\*
- 5: /cgn2\_6/prodata/1/pubppa/US10B\_PUBCOMB.pep:\*
- 6: /cgn2\_6/prodata/1/pubppa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	3	US-09-999-745-67
2	1214	100.0	225	3	US-09-866-538-12
3	1214	100.0	225	3	US-09-794-308-12
4	1214	100.0	225	3	US-09-865-291-12
5	1214	100.0	225	4	US-10-006-922-12
6	1214	100.0	225	4	US-10-081-864-8
7	1214	100.0	225	4	US-10-121-358-1
8	1214	100.0	225	4	US-10-135-920-2
9	1214	100.0	225	4	US-10-132-067-4
10	1214	100.0	225	4	US-10-370-570-56
11	1214	100.0	225	4	US-10-406-618-32
12	1214	100.0	225	4	US-10-433-640-13
13	1214	100.0	225	4	US-10-311-030-7
14	1214	100.0	225	4	US-10-845-484-3
15	1214	100.0	225	5	US-10-885-988-12
16	1214	100.0	225	5	US-10-656-029-22
17	1214	100.0	225	5	US-10-857-622-12
18	1214	100.0	225	5	US-10-505-486-26
19	1214	100.0	225	5	US-10-844-064A-2
20	1214	100.0	225	5	US-10-931-304-1
21	1214	100.0	487	5	US-10-343-977-1
22	1214	100.0	487	5	US-10-343-977-1
23	1214	100.0	506	4	US-10-343-977-2
24	1214	100.0	506	4	US-10-343-977-2
25	1214	100.0	547	5	US-10-343-977-3
26	1214	100.0	547	5	US-10-343-977-3
27	1210	99.7	226	4	US-10-311-030-9

28	1210	99.7	240	4	US-10-152-296-2	Sequence 2, Appl1
29	1210	99.7	240	4	US-10-739-656-2	Sequence 2, Appl1
30	1210	99.7	548	4	US-10-214-932-52	Sequence 52, Appl1
31	1210	99.7	548	4	US-10-214-932-76	Sequence 76, Appl1
32	1207	99.4	225	4	US-10-121-258-20	Sequence 20, Appl1
33	1207	99.4	225	4	US-10-315-920-4	Sequence 4, Appl1
34	1207	99.4	225	5	US-10-931-304-20	Sequence 20, Appl1
35	1196	98.5	236	4	US-10-314-936-2	Sequence 2, Appl1
36	1196	98.5	236	4	US-10-314-936-4	Sequence 4, Appl1
37	1196	98.5	236	6	US-11-021-014-2	Sequence 2, Appl1
38	1196	98.5	236	6	US-11-021-014-4	Sequence 4, Appl1
39	1195.5	98.5	226	4	US-10-724-178-12	Sequence 12, Appl1
40	1191	98.1	225	4	US-10-315-920-6	Sequence 6, Appl1
41	1186	97.7	225	4	US-10-442-148A-7	Sequence 7, Appl1
42	1186	97.7	239	4	US-10-442-148A-8	Sequence 8, Appl1
43	1160	95.6	225	4	US-10-121-258-4	Sequence 4, Appl1
44	1160	95.6	225	5	US-10-931-304-4	Sequence 4, Appl1
45	1153	95.0	225	4	US-10-121-258-24	Sequence 24, Appl1

## ALIGNMENTS

```
RESULT 1
US-09-999-745-67
; Sequence 67, Application US/09999745
; Patent No. US20020157120A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger Y.
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: REG1470-1
; CURRENT APPLICATION NUMBER: US/09/999,745
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/316,920
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
; LENGTH: 225
; TYPE: PRP
; ORGANISM: Diacrosoma sp.
US-09-999-745-67

Query Match      100.0%; Score 1214; DB 3; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVMEGTUNGHEFEIRGEGRPYEGHNTVKLKVTGGFLPFAWDI 60
        |||
DB      1 MRSSKNVKEFMRFKVMEGTUNGHEFEIRGEGRPYEGHNTVKLKVTGGFLPFAWDI 60

QY      61 LSPQFOYGSVYVYKHPADIPDYKKLSFPEGFKMERVWNPFDGCVVTVTQDSIADGCFIY 120
        |||
DB      61 LSPQFOYGSVYVYKHPADIPDYKKLSFPEGFKMERVWNPFDGCVVTVTQDSIADGCFIY 120

QY      121 KYVFIGNPSPDPVWQKKTWGEASTERLYPRDGVYKGEIHKALKIKDGGHYLVEFKSI 180
        |||
DB      121 KYVFIGNPSPDPVWQKKTWGEASTERLYPRDGVYKGEIHKALKIKDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYVYVDSKDLTSHNEDYTVVEQYERTGHHFL 225
        |||
DB      181 YMAKKPVQLPGYVYVDSKDLTSHNEDYTVVEQYERTGHHFL 225

RESULT 2
US-09-866-538-12
; Sequence 12, Application US/09866538
; Publication No. US20030032088A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger
```

```

; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530-2
; CURRENT APPLICATION NUMBER: US/09/866,538
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-866-538-12

Query Match          100.0%; Score 1214; DB 3; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
1 MRSSKNVTKKEFMRFKVRMEGTGNGHGFELIEGEGRPYEGHNTVTKLKYTKGGPLPFANDI 60
|
|
|
61 LSPQOYSGSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTVQDSSLQDGCFTY 120
|
|
|
61 KYKFGVNPSPDGPVMOKKTGMWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
|
|
|
121 KYKFGVNPSPDGPVMOKKTGMWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
|
|
|
121 KYKFGVNPSPDGPVMOKKTGMWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180

QY
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

Db
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

QY
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

Db
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 3
US-09-794-308-12
; Sequence 12, Application US/09794308
; Publication No. US20030170911A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSIEN, Roger
; APPLICANT: ZACHARIAS, David
; APPLICANT: BAIRD, Geoffrey
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530
; CURRENT APPLICATION NUMBER: US/09/794,308
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-794-308-12

Query Match          100.0%; Score 1214; DB 3; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
1 MRSSKNVTKKEFMRFKVRMEGTGNGHGFELIEGEGRPYEGHNTVTKLKYTKGGPLPFANDI 60
|
|
|
61 LSPQOYSGSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTVQDSSLQDGCFTY 120
|
|
|
61 LSPQOYSGSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTVQDSSLQDGCFTY 120
|
|
|
61 KYKFGVNPSPDGPVMOKKTGMWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
|
|
|
121 KYKFGVNPSPDGPVMOKKTGMWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
|
|
|
121 KYKFGVNPSPDGPVMOKKTGMWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180

QY
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

Db
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

QY
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

Db
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

```

```

Db
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 4
US-09-865-291-12
; Sequence 12, Application US/09865291
; Publication No. US20030186229A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSIEN, Roger
; APPLICANT: TING, Alice
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-865-291-12

Query Match          100.0%; Score 1214; DB 3; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db
1 MRSSKNVTKKEFMRFKVRMEGTGNGHGFELIEGEGRPYEGHNTVTKLKYTKGGPLPFANDI 60
|
|
|
61 LSPQOYSGSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTVQDSSLQDGCFTY 120
|
|
|
61 LSPQOYSGSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTVQDSSLQDGCFTY 120
|
|
|
61 KYKFGVNPSPDGPVMOKKTGMWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
|
|
|
121 KYKFGVNPSPDGPVMOKKTGMWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
|
|
|
121 KYKFGVNPSPDGPVMOKKTGMWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180

QY
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

Db
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

QY
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

Db
181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 5
US-10-006-922-12
; Sequence 12, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A.
; APPLICANT: Pradkov, Arcady F.
; APPLICANT: Ladas, Yulii A.
; APPLICANT: Matcz, Mikhail V.
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and
; FILE REFERENCE: CLON-035C1P
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46

```



```
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Discoecia species
US-10-006-922-12
```

```
Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRSSKNVIEKPMRFKRMGTGNGHFEIIEGEGGRPYEGHNTVKLKVTGGLPFPAMD 60
    |||||
DB 1 MRSSKNVIEKPMRFKRMGTGNGHFEIIEGEGGRPYEGHNTVKLKVTGGLPFPAMD 60
QY 61 LSPQFGSKVYVKGHPADIPDYKGLSPFGFKMERVMNPFDDGVTVTODSLQDGCFTY 120
    |||||
DB 61 LSPQFGSKVYVKGHPADIPDYKGLSPFGFKMERVMNPFDDGVTVTODSLQDGCFTY 120
QY 121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGYLKGEIHKALKKDGSHYLVFKSI 180
    |||||
DB 121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGYLKGEIHKALKKDGSHYLVFKSI 180
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
    |||||
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 6
US-10-081-864-8
/ Sequence 8, Application US/10081864
/ Publication No. US2003022287A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Lukanov, Sergey
/ APPLICANT: Lukanov, Konstantin
/ APPLICANT: Yanushevich, Yuriy
/ APPLICANT: Savitsky, Alexander
/ APPLICANT: Pradkov, Arcady
/ TITLE OF INVENTION: No. US2003022287A1 Aggregating Fluorescent Proteins and
/ TITLE OF INVENTION: Methode for Using the Same
/ FILE REFERENCE: CLON-067
/ CURRENT APPLICATION NUMBER: US/10/081,864
/ PRIOR FILING DATE: 2002-06-19
/ PRIOR APPLICATION NUMBER: 10/006,922
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 60/270,983
/ PRIOR FILING DATE: 2001-02-21
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 8
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Discoecia sp.
US-10-081-864-8
```

```
Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRSSKNVIEKPMRFKRMGTGNGHFEIIEGEGGRPYEGHNTVKLKVTGGLPFPAMD 60
    |||||
DB 1 MRSSKNVIEKPMRFKRMGTGNGHFEIIEGEGGRPYEGHNTVKLKVTGGLPFPAMD 60
QY 61 LSPQFGSKVYVKGHPADIPDYKGLSPFGFKMERVMNPFDDGVTVTODSLQDGCFTY 120
    |||||
DB 61 LSPQFGSKVYVKGHPADIPDYKGLSPFGFKMERVMNPFDDGVTVTODSLQDGCFTY 120
QY 121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGYLKGEIHKALKKDGSHYLVFKSI 180
    |||||
DB 121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGYLKGEIHKALKKDGSHYLVFKSI 180
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
    |||||
```

```
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 7
US-10-121-258-1
/ Sequence 1, Application US/10121258
/ Publication No. US20030059835A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Tsien, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: UC083,1CP2CPI
/ CURRENT APPLICATION NUMBER: US/10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Discoecia sp.
/ FEATURE: misc_feature
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(225)
/ OTHER INFORMATION: wild-type DsRed
US-10-121-258-1
```

```
Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRSSKNVIEKPMRFKRMGTGNGHFEIIEGEGGRPYEGHNTVKLKVTGGLPFPAMD 60
    |||||
DB 1 MRSSKNVIEKPMRFKRMGTGNGHFEIIEGEGGRPYEGHNTVKLKVTGGLPFPAMD 60
QY 61 LSPQFGSKVYVKGHPADIPDYKGLSPFGFKMERVMNPFDDGVTVTODSLQDGCFTY 120
    |||||
DB 61 LSPQFGSKVYVKGHPADIPDYKGLSPFGFKMERVMNPFDDGVTVTODSLQDGCFTY 120
QY 121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGYLKGEIHKALKKDGSHYLVFKSI 180
    |||||
DB 121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGYLKGEIHKALKKDGSHYLVFKSI 180
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
    |||||
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 8
US-10-315-920-2
```

```
/ Sequence 2, Application US/10315920
/ Publication No. US20030175809A1
/ GENERAL INFORMATION:
/ APPLICANT: Pradkov, Arcady Pedorovich
/ APPLICANT: Tersikh, Alexey
/ TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
/ TITLE OF INVENTION: FOR THEIR USE
/ FILE REFERENCE: CLON-077CIP
/ CURRENT APPLICATION NUMBER: US/10/315,920
/ PRIOR FILING DATE: 2002-12-09
/ PRIOR APPLICATION NUMBER: 60/211,607
/ PRIOR FILING DATE: 2000-06-14
/ PRIOR APPLICATION NUMBER: PCT/US01/19097
/ PRIOR FILING DATE: 2001-06-13
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 225
/ TYPE: PRT
```

ORGANISM: Discosoma sp.  
US-10-315-920-2

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6,6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
| MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
DB 1 MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
| MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
QY 61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
| LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
| LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
QY 121 KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
| KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
| KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
| YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
| YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

## RESULT 9

US-10-132-067-4

Sequence 4, Application US/10132067  
Publication No. US20030203355A1  
GENERAL INFORMATION:  
APPLICANT: Bradbury, Andrew  
APPLICANT: Zeytun, Ahmet  
APPLICANT: Waldo, Geoffrey  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Fluorobodies: Binding Ligands With Intrinsic  
TITLE OF INVENTION: Fluorescence  
FILE REFERENCE: 021362-00060005  
CURRENT APPLICATION NUMBER: US/10/132,067  
CURRENT FILING DATE: 2002-04-24  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
FEATURE:  
OTHER INFORMATION: red fluorescent protein (dared)  
US-10-132-067-4

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6,6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
| MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
DB 1 MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
| MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
QY 61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
| LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
| LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
QY 121 KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
| KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
| KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
| YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
| YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

## RESULT 10

US-10-370-570-56

Sequence 56, Application US/10370570  
Publication No. US20030219717A1

GENERAL INFORMATION:  
APPLICANT: DAHL, Soren Weis et al.  
TITLE OF INVENTION: FLUOROPHORE COMPLEMENTATION PRODUCTS  
FILE REFERENCE: 3759-0130P  
CURRENT APPLICATION NUMBER: US/10/370,570  
CURRENT FILING DATE: 2003-02-24  
NUMBER OF SEQ ID NOS: 71  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 56  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
US-10-370-570-56

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6,6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
| MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
DB 1 MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
| MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
QY 61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
| LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
| LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
QY 121 KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
| KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
| KVFPGVNFPSDGPVWOKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
| YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
| YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

## RESULT 11

US-10-406-618-32

Sequence 32, Application US/10406618  
Publication No. US20030219814A1  
GENERAL INFORMATION:  
APPLICANT: Wan, David Chi-Cheong  
APPLICANT: Ip, Dennis Tsz-Ming  
APPLICANT: The Chinese University of Hong Kong  
TITLE OF INVENTION: No. US20030219814A1el Fluorescent Proteins  
FILE REFERENCE: 016285-34-1  
CURRENT APPLICATION NUMBER: US/10/406,618  
CURRENT FILING DATE: 2003-04-02  
PRIOR APPLICATION NUMBER: US 60/387,968  
PRIOR FILING DATE: 2002-06-11  
PRIOR APPLICATION NUMBER: US 60/370,598  
PRIOR FILING DATE: 2002-04-05  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 32  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE: Description of Artificial Sequence: Spontaneously  
OTHER INFORMATION: Fluorescent protein Dared.  
US-10-406-618-32

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6,6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
| MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
DB 1 MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
| MRSSKNVKEFMRFKVRMEGTNGHHEFEIEGEGRPYEGHNTVYLKVTKGGLPFPAMD 60  
QY 61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
| LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120  
| LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120

```
QY 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
DB 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
QY 181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225
DB 181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 12
US-10-433-640-13
; Sequence 13, Application US/10433640
; Publication No. US20040115792A1
; GENERAL INFORMATION:
; APPLICANT: Lachenberg-Fraete, Heila
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF
; FILE REFERENCE: 1487/3
; CURRENT APPLICATION NUMBER: US/10/433,640
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/14610
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: DE 10061872.3
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Dicosoma sp.
US-10-433-640-13

Query Match
Best Local Similarity 100.0%; Score 1214; DB 4; Length 225;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSKVNIKEFMKFKYRMESTVNGHBEFIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60
DB 1 MSSSKVNIKEFMKFKYRMESTVNGHBEFIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60
QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERNVNFDDGVVYVTTODSSLODGCFTY 120
DB 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERNVNFDDGVVYVTTODSSLODGCFTY 120
QY 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
DB 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
QY 181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225
DB 181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 13
US-10-311-030-7
; Sequence 7, Application US/10311030
; Publication No. US20040171107A1
; GENERAL INFORMATION:
; APPLICANT: Nelson, David
; APPLICANT: Zamiatra, Elize
; APPLICANT: Telen, Roger
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS
; FILE REFERENCE: 15916-032US1
; CURRENT APPLICATION NUMBER: US/10/311,030
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US01/04625
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US 60/184,732
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
```

```
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-311-030-7

Query Match
Best Local Similarity 100.0%; Score 1214; DB 4; Length 225;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSKVNIKEFMKFKYRMESTVNGHBEFIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60
DB 1 MSSSKVNIKEFMKFKYRMESTVNGHBEFIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60
QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERNVNFDDGVVYVTTODSSLODGCFTY 120
DB 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERNVNFDDGVVYVTTODSSLODGCFTY 120
QY 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
DB 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
QY 181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225
DB 181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 14
US-10-845-484-3
; Sequence 3, Application US/10845484
; Publication No. US20040248180A1
; GENERAL INFORMATION:
; APPLICANT: Bulina, Maria
; APPLICANT: Chudakov, Dmitry
; APPLICANT: Lukyanov, Konstantin
; TITLE OF INVENTION: Mutant Chromophores/Fluorophores and
; FILE REFERENCE: CLOM 092
; CURRENT APPLICATION NUMBER: US/10/845,484
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/343128
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/US02/41418
; PRIOR FILING DATE: 2002-12-23
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Dicosoma sp.
US-10-845-484-3

Query Match
Best Local Similarity 100.0%; Score 1214; DB 5; Length 225;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSKVNIKEFMKFKYRMESTVNGHBEFIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60
DB 1 MSSSKVNIKEFMKFKYRMESTVNGHBEFIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60
QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERNVNFDDGVVYVTTODSSLODGCFTY 120
DB 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERNVNFDDGVVYVTTODSSLODGCFTY 120
QY 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
DB 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
QY 181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225
DB 181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225
```

```
RESULT 15
US-10-885-988-12
; Sequence 12, Application US/10885988
; Publication No. US20040259165A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530-2
; CURRENT APPLICATION NUMBER: US/10/885,988
; CURRENT FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: US/09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discoosoma sp.
; US-10-885-988-12

Query Match      100.0%; Score 1214; DB 5; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVTKERFRKVRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTGKGPLPAMD I 60
        |||||||
DB      1 MRSSKNVTKERFRKVRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTGKGPLPAMD I 60
        |||||||

QY      61 LSPQFQYSGKYVVKHPADI PDYKKLSFPRGFKERVMNPFEDGVVTVTQDSSLQDGCFTY 120
        |||||||
DB      61 LSPQFQYSGKYVVKHPADI PDYKKLSFPRGFKERVMNPFEDGVVTVTQDSSLQDGCFTY 120
        |||||||

QY      121 KVFPIGVNFPSPDGPVMQKKTGMGEASTERLYPRDGVLKGEIHKALKLKDGGHYLVFPKSI 180
        |||||||
DB      121 KVFPIGVNFPSPDGPVMQKKTGMGEASTERLYPRDGVLKGEIHKALKLKDGGHYLVFPKSI 180
        |||||||

QY      181 YMAKPPVQLPGYYVYDSKLDITSNEDYTIYEQYERTEGRHHLFL 225
        |||||||
DB      181 YMAKPPVQLPGYYVYDSKLDITSNEDYTIYEQYERTEGRHHLFL 225
        |||||||
```

Search completed: January 11, 2006, 02:06:37  
Job time : 63 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 11, 2006, 02:03:07 ; Search time 8 Seconds  
(without alignments)  
238.884 Million cell updates/sec

Title: US-10-006-922a-12

Perfect score: 1214  
Sequence: 1 MRSSKVIKFKFMRFKRMKG.....EDTYVEQYERTGHHFL 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 61141 seqs, 8493638 residues

Total number of hits satisfying chosen parameters: 61141

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : PubliSearch Applications\_AA\_New:\*  
1: /cgn2\_6/prodata/1/pubppa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/prodata/1/pubppa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/1/pubppa/US07\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/1/pubppa/PCT\_NEW\_PUB.pep:\*  
5: /cgn2\_6/prodata/1/pubppa/US09\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/1/pubppa/US10\_NEW\_PUB.pep:\*  
7: /cgn2\_6/prodata/1/pubppa/US11\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/1/pubppa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	100.0	225	6	US-10-209-208-1
2	1207	99.4	225	6	US-10-209-208-20
3	1160	95.6	225	6	US-10-209-208-4
4	1153	95.0	225	6	US-10-209-208-24
5	1121	92.3	226	6	US-10-209-208-6
6	1021	84.1	225	6	US-10-209-208-8
7	1012	83.4	225	6	US-10-209-208-79
8	485.5	40.0	238	7	US-11-179-411-16
9	485.5	40.0	238	7	US-11-179-766-16
10	473.5	39.0	238	7	US-11-179-411-32
11	473.5	39.0	238	7	US-11-175-766-32
12	422.5	34.8	233	7	US-11-179-411-27
13	422.5	34.8	233	7	US-11-175-766-27
14	410.5	33.8	233	7	US-11-179-411-33
15	410.5	33.8	233	7	US-11-175-766-33
16	210.5	17.3	233	6	US-10-209-208-10
17	210.5	17.3	233	7	US-11-089-551A-40
18	209.5	17.3	239	6	US-10-209-208-12
19	205.5	16.9	239	6	US-10-209-208-15
20	204.5	16.8	512	7	US-11-032-236-6
21	204.5	16.8	1219	7	US-11-032-236-4
22	200.5	15.5	239	6	US-10-209-208-13
23	192.5	15.9	239	6	US-10-209-208-11
24	192.5	15.9	239	6	US-10-209-208-14
25	183.5	15.1	437	7	US-11-069-642-92

26	182.5	15.0	438	7	US-11-069-642-47	Sequence 47, App1
27	182.5	15.0	438	7	US-11-069-642-49	Sequence 49, App1
28	182.5	15.0	438	7	US-11-069-642-51	Sequence 51, App1
29	182.5	15.0	438	7	US-11-069-642-53	Sequence 53, App1
30	182.5	15.0	438	7	US-11-069-642-55	Sequence 55, App1
31	182.5	15.0	438	7	US-11-069-642-57	Sequence 57, App1
32	182.5	15.0	438	7	US-11-069-642-59	Sequence 59, App1
33	182.5	15.0	438	7	US-11-069-642-61	Sequence 61, App1
34	182.5	15.0	438	7	US-11-069-642-63	Sequence 63, App1
35	103	8.5	559	6	US-10-793-626-1376	Sequence 1376, App
36	80	6.6	424	6	US-10-467-657-90	Sequence 90, App1
37	80	6.6	424	6	US-10-467-657-90	Sequence 90, App1
38	80	6.6	690	7	US-11-052-554A-232	Sequence 232, App
39	76.5	6.3	325	6	US-10-467-657-14	Sequence 14, App1
40	76.5	6.3	325	6	US-10-467-657-5568	Sequence 5568, App
41	76.5	6.3	574	7	US-11-053-100-40	Sequence 40, App1
42	76.5	6.3	724	7	US-11-053-100-41	Sequence 41, App1
43	76.5	6.3	874	7	US-11-053-100-42	Sequence 42, App1
44	76.5	6.3	1174	7	US-11-053-100-43	Sequence 43, App1
45	76.5	6.3	1663	6	US-10-055-877-148	Sequence 148, App

## ALIGNMENTS

```
RESULT 1
US-10-209-208-1
; Sequence 1, Application US/10209208
; Publication No. US20050244921A1
GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Drosophila sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(225)
; OTHER INFORMATION: wild-type DsRed
US-10-209-208-1
Query Match 100.0%; Score 1214; DB 6; Length 225;
Best Local Similarity 100.0%; Pred. No. 2.3e-110;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 60
1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 60
Db 1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 60
QY 1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 120
1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 120
Db 1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 120
QY 1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 180
1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 180
Db 1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 180
QY 1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 225
1 MRSSKVIKFKFMRFKRMKGHEPELIGBEGRGYBGNVVKLKVTKGGLPFWMDI 225
```

```
Db      181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
|||||
RESULT 2
US-10-209-208-20
; Sequence 20, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dated with 1125R
US-10-209-208-20

Query Match      99.4%; Score 1207; DB 6; Length 225;
Best Local Similarity 99.6%; Pred. No. 1.1e-109;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db      1 MSSKNVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
1 MSSKNVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
Qy      61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
Db      61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
Qy      121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
Db      121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
Qy      181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
Db      181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225

RESULT 3
US-10-209-208-4
; Sequence 4, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
```

```
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dated polypeptide variant "T1"
US-10-209-208-4

Query Match      95.6%; Score 1160; DB 6; Length 225;
Best Local Similarity 96.0%; Pred. No. 3.7e-105;
Matches 216; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy      1 MSSKNVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
1 MSSKNVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
Db      1 MASSEDVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
1 MASSEDVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
Qy      61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
Db      61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
Qy      121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
Db      121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
Qy      181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
Db      181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225

RESULT 4
US-10-209-208-24
; Sequence 24, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dated polypeptide variant "T1" with 1125R mutation
US-10-209-208-24

Query Match      95.0%; Score 1153; DB 6; Length 225;
Best Local Similarity 95.6%; Pred. No. 1.8e-104;
Matches 215; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy      1 MSSKNVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
1 MSSKNVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
Db      1 MASSEDVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
1 MASSEDVKEFMRPFKVRMEGTVNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFANDI 60
Qy      61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
Db      61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
61 LSPQFQSGSKYVVKHPADIPDYKLSFPEGFKMERVMNFBEDGVVTVTQDSSLQDGCFTY 120
Qy      121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
Db      121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
121 KYKFGVNFPSDGPVWQKKTWGMWASTERTLYPRDGVLTGKEIHKALKLKDGGHYLVFPKSI 180
Qy      181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
Db      181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
181 YMAKRPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225
```

Db 121 KYKFGVNPSPGPMQKKTWMEASTELYPKIDVGLKGEIHKALKDGGHYLVEFKSI 180  
 QY 181 YMAKPPVOLPGYIYVDSKIDITSHNEDYTIIVEQYERABGRHHLFL 225  
 Db 181 YMAKPPVOLPGYIYVDSKIDITSHNEDYTIIVEQYERABGRHHLFL 225

RESULT 5  
 US-10-209-208-6  
 / Sequence 6, Application US/10209208  
 / Publication No. US20050244921A1  
 / GENERAL INFORMATION:

APPLICANT: Telen, Roger  
 APPLICANT: Campbell, Robert  
 APPLICANT: Geoffrey Baird  
 TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
 TITLE OF INVENTION: FOR MAKING SAME  
 FILE REFERENCE: UC083.1CP2CP2  
 CURRENT APPLICATION NUMBER: US/10/209,208  
 CURRENT FILING DATE: 2002-07-29  
 PRIOR APPLICATION NUMBER: 10/121,258  
 PRIOR FILING DATE: 2002-04-10  
 PRIOR APPLICATION NUMBER: 09/866,538  
 PRIOR FILING DATE: 2001-05-24  
 PRIOR APPLICATION NUMBER: 09/794,308  
 PRIOR FILING DATE: 2001-02-26  
 NUMBER OF SEQ ID NOS: 80  
 SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6  
 LENGTH: 226  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Dated polypeptide variant "dimer2"

US-10-209-208-6

Query Match 92.3%; Score 1121; DB 6; Length 226;  
 Best Local Similarity 92.8%; Pred. No. 2,2e-101;  
 Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 3 SSKNVIKKFMKFRVYRMNGTVNGHGFELISGEGRPYEGHNTVTLKTKGGLPFPAMDILS 62  
 Db 4 SSSDVLIKEFMKFRVYRMNGTVNGHGFELISGEGRPYEGHNTVTLKTKGGLPFPAMDILS 63  
 QY 63 POFQYGSKAVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFTIYK 122  
 Db 64 POFQYGSKAVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFTIYK 123  
 QY 123 KPIGVNFPSPGPMQKKTWMEASTELYPKIDVGLKGEIHKALKDGGHYLVEFKSIY 182  
 Db 124 KPIGVNFPSPGPMQKKTWMEASTELYPKIDVGLKGEIHKALKDGGHYLVEFKSIY 183  
 QY 183 AKKPPVOLPGYIYVDSKIDITSHNEDYTIIVEQYERABGRHHLFL 225  
 Db 184 AKKPPVOLPGYIYVDSKIDITSHNEDYTIIVEQYERABGRHHLFL 226

RESULT 6  
 US-10-209-208-8  
 / Sequence 8, Application US/10209208  
 / Publication No. US20050244921A1  
 / GENERAL INFORMATION:  
 APPLICANT: Telen, Roger  
 APPLICANT: Campbell, Robert  
 APPLICANT: Geoffrey Baird  
 TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
 TITLE OF INVENTION: FOR MAKING SAME  
 FILE REFERENCE: UC083.1CP2CP2  
 CURRENT APPLICATION NUMBER: US/10/209,208  
 CURRENT FILING DATE: 2002-07-29  
 PRIOR APPLICATION NUMBER: 10/121,258  
 PRIOR FILING DATE: 2002-04-10

PRIOR APPLICATION NUMBER: 09/866,538  
 PRIOR FILING DATE: 2001-05-24  
 PRIOR APPLICATION NUMBER: 09/794,308  
 PRIOR FILING DATE: 2001-02-26  
 NUMBER OF SEQ ID NOS: 80  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 8  
 LENGTH: 225  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Dated polypeptide variant "mrFP1"

US-10-209-208-8

Query Match 84.1%; Score 1021; DB 6; Length 225;  
 Best Local Similarity 86.9%; Pred. No. 1e-91;  
 Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSSKNVIEKFMKFRVYRMNGTVNGHGFELISGEGRPYEGHNTVTLKTKGGLPFPAMD 60  
 Db 1 MSSSDVLIKEFMKFRVYRMNGTVNGHGFELISGEGRPYEGHNTVTLKTKGGLPFPAMD 60  
 QY 61 LSPQYGSKAVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFTIY 120  
 Db 61 LSPQYGSKAVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFTIY 120  
 QY 121 KYKFGVNPSPGPMQKKTWMEASTELYPKIDVGLKGEIHKALKDGGHYLVEFKSI 180  
 Db 121 KYKFGVNPSPGPMQKKTWMEASTELYPKIDVGLKGEIHKALKDGGHYLVEFKSI 180  
 QY 181 YMAKPPVOLPGYIYVDSKIDITSHNEDYTIIVEQYERABGRH 221  
 Db 181 YMAKPPVOLPGYIYVDSKIDITSHNEDYTIIVEQYERABGRH 221

RESULT 7  
 US-10-209-208-79

/ Sequence 79, Application US/10209208  
 / Publication No. US20050244921A1  
 / GENERAL INFORMATION:

APPLICANT: Telen, Roger  
 APPLICANT: Campbell, Robert  
 APPLICANT: Geoffrey Baird  
 TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
 TITLE OF INVENTION: FOR MAKING SAME  
 FILE REFERENCE: UC083.1CP2CP2  
 CURRENT APPLICATION NUMBER: US/10/209,208  
 CURRENT FILING DATE: 2002-07-29  
 PRIOR APPLICATION NUMBER: 10/121,258  
 PRIOR FILING DATE: 2002-04-10  
 PRIOR APPLICATION NUMBER: 09/866,538  
 PRIOR FILING DATE: 2001-05-24  
 PRIOR APPLICATION NUMBER: 09/794,308  
 PRIOR FILING DATE: 2001-02-26  
 NUMBER OF SEQ ID NOS: 80  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 79  
 LENGTH: 225  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Dated polypeptide variant "mrFP1.1"

US-10-209-208-79

Query Match 83.4%; Score 1012; DB 6; Length 225;  
 Best Local Similarity 86.0%; Pred. No. 7,4e-91;  
 Matches 190; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 1 MSSKNVIEKFMKFRVYRMNGTVNGHGFELISGEGRPYEGHNTVTLKTKGGLPFPAMD 60  
 Db 1 MSSSDVLIKEFMKFRVYRMNGTVNGHGFELISGEGRPYEGHNTVTLKTKGGLPFPAMD 60  
 QY 61 LSPQYGSKAVYKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFTIY 120

Db 61 LSPQWMSKAVVKGADIPDYKLKLSFPGFKMEKVMNPFEDGGVTVTQDSSLQDGEFLY 120  
Qy 121 KKRFIGNPSPGPMQKKTMGMEASTERLYPRDGLKEIHAKLKADGGHYLVEFSI 180  
Db 121 KKRLGTFPPSGPMQKKTMGMEASERMPEDALKEIMRLKLDGGHYDAEVKTT 180  
Qy 181 YMAKKPVOLPGYVYVDSKLDITSNEDYTYVEQYRTEGRH 221  
Db 181 YMAKKPVOLPGAYKTDIKLDITSNEDITYVEQYRABGRH 221

RESULT 8  
US-11-179-411-16  
; Sequence 16, Application US/11179411  
; Publication No. US20050266491A1  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: Szczepaniak, William  
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING T  
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH  
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
; FILE REFERENCE: 24729-128  
; CURRENT APPLICATION NUMBER: US/11/179,411  
; PRIOR FILING DATE: 2005-07-12  
; PRIOR APPLICATION NUMBER: US/09/808,898  
; PRIOR FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/189,691  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 09/277,716  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 08/757,046  
; PRIOR FILING DATE: 1996-11-25  
; PRIOR APPLICATION NUMBER: 08/597,274  
; PRIOR FILING DATE: 1996-02-06  
; PRIOR APPLICATION NUMBER: 08/908,909  
; PRIOR FILING DATE: 1997-08-08  
; PRIOR APPLICATION NUMBER: 08/990,103  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Renilla mulleri  
US-11-179-411-16

Query Match 40.0%; Score 485.5; DB 7; Length 238;  
Best Local Similarity 44.5%; Pred. No. 5.8e-40;  
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;  
Qy 4 SKNVIK-----EFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTYKGGLPPAW 58  
Db 2 SKQILKNTCLQEWMSYKVNLEGIIVNNHVFTEMGCKGNILFGNQLVQIRVTKGAPLPFAF 61  
Qy 59 DILSPQWMSKAVVKGADIPDYKLKLSFPGFKMEKVMNPFEDGGVTVTQDSSLQDGEFLY 118  
Db 62 DIVSPAFQYGNRTFTKYPNDISDYFIQSPFAGFYMERLTLYEDGGLVEIRSDINLIDKX 121  
Qy 119 IYKVFIGNPSPGPMQKKTMGMEASTERLYPRDGLKEIHAKLKADGGHYLVEFSI 178  
Db 122 YRVEYKSNFPPDGGPMQKKTILGIEPSFEAMYNNGVLVGEVILVYKLNSGKTYSCMK 181  
Qy 179 SIYMAKKPV-OLPGYVYVDSKLDITSNEDYTYVEQYRTEGRH 215  
Db 182 TLMKSKGVVKEFPSPYHFIQRLKLT-YVEDGGFVEQHE 218

RESULT 9  
US-11-175-766-16  
; Sequence 16, Application US/11175766  
; Publication No. US2005027211A1

; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: Szczepaniak, William  
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING T  
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH  
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
; FILE REFERENCE: 24729-128  
; CURRENT APPLICATION NUMBER: US/11/175,766  
; PRIOR FILING DATE: 2005-07-06  
; PRIOR APPLICATION NUMBER: US/09/808,898  
; PRIOR FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/189,691  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 09/277,716  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 08/757,046  
; PRIOR FILING DATE: 1996-11-25  
; PRIOR APPLICATION NUMBER: 08/597,274  
; PRIOR FILING DATE: 1996-02-06  
; PRIOR APPLICATION NUMBER: 08/908,909  
; PRIOR FILING DATE: 1997-08-08  
; PRIOR APPLICATION NUMBER: 08/990,103  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Renilla mulleri  
US-11-175-766-16

Query Match 40.0%; Score 485.5; DB 7; Length 238;  
Best Local Similarity 44.5%; Pred. No. 5.8e-40;  
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;  
Qy 4 SKNVIK-----EFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTYKGGLPPAW 58  
Db 2 SKQILKNTCLQEWMSYKVNLEGIIVNNHVFTEMGCKGNILFGNQLVQIRVTKGAPLPFAF 61  
Qy 59 DILSPQWMSKAVVKGADIPDYKLKLSFPGFKMEKVMNPFEDGGVTVTQDSSLQDGEFLY 118  
Db 62 DIVSPAFQYGNRTFTKYPNDISDYFIQSPFAGFYMERLTLYEDGGLVEIRSDINLIDKX 121  
Qy 119 IYKVFIGNPSPGPMQKKTMGMEASTERLYPRDGLKEIHAKLKADGGHYLVEFSI 178  
Db 122 YRVEYKSNFPPDGGPMQKKTILGIEPSFEAMYNNGVLVGEVILVYKLNSGKTYSCMK 181  
Qy 179 SIYMAKKPV-OLPGYVYVDSKLDITSNEDYTYVEQYRTEGRH 215  
Db 182 TLMKSKGVVKEFPSPYHFIQRLKLT-YVEDGGFVEQHE 218

RESULT 10  
US-11-179-411-32  
; Sequence 32, Application US/11179411  
; Publication No. US20050266491A1  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: Szczepaniak, William  
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING T  
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH  
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
; FILE REFERENCE: 24729-128  
; CURRENT APPLICATION NUMBER: US/11/179,411  
; PRIOR FILING DATE: 2005-07-12  
; PRIOR APPLICATION NUMBER: US/09/808,898  
; PRIOR FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/189,691  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 09/277,716  
; PRIOR FILING DATE: 1999-03-26



PRIOR APPLICATION NUMBER: 08/757,046  
 PRIOR FILING DATE: 1996-11-25  
 PRIOR APPLICATION NUMBER: 08/597,274  
 PRIOR FILING DATE: 1996-02-06  
 PRIOR APPLICATION NUMBER: 08/908,909  
 PRIOR FILING DATE: 1997-08-08  
 PRIOR APPLICATION NUMBER: 08/990,103  
 PRIOR FILING DATE: 1997-12-12  
 NUMBER OF SEQ ID NOS: 33  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 32  
 LENGTH: 238  
 TYPE: PRF  
 ORGANISM: *Ptilosarcus gurneyi*  
 US-11-179-411-32

Query Match 39.0%; Score 473.5; DB 7; Length 238;  
 Best Local Similarity 45.5%; Pred. No. 8.3e-39;  
 Matches 97; Conservative 40; Mismatches 73; Indels 3; Gaps 3;

5 KNV-IKEFMRFKVMEGTVNGHEFEIEGEGRPRYEGHNTVKLKTKGGLPFAMDILSP 63  
 7 KNTGLKEIMSAKASVSGIYNNHVFMEGFGKGNVLFGNQLMQRVTYKGGPLPFADIVSI 66  
 64 OFQYSGKYVVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTODSSLQDGCFTIYKYK 123  
 67 AFQYGNRTFTKYDDIADYFVQSPFAGFYERNLRFEDGALVDIRSDISLEDKCFHYKYE 126  
 124 FIGVNPSPDGPVWQKKTGMGEASTELYPKDGVLKGEIHKALKLKDGGHYLVEFKSIYMA 183  
 127 YRNGPSPNGPVWQKAILGMEPSFEVYVWNSGVLVGEVDLVYKLSGNTYSCHMKTFYRS 186  
 184 KKPV-QLPGYVYVDSKLDITSHNEDYTIYEQYE 215  
 187 KGGVKEPPEYHFIHRLKLT-YVEBGSFVEQHE 218

RESULT 11  
 US-11-175-766-32  
 Sequence 32, Application US/11175766  
 Publication No. US2005027211A1

GENERAL INFORMATION:  
 APPLICANT: Bryan, Bruce  
 APPLICANT: Szent-Gyorgyi, Christopher  
 APPLICANT: Szczeniak, William  
 TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING T  
 TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
 FILE REFERENCE: 24729-128  
 CURRENT APPLICATION NUMBER: US/11/175,766  
 CURRENT FILING DATE: 2005-07-06  
 PRIOR APPLICATION NUMBER: US/09/808,898  
 PRIOR FILING DATE: 2001-03-15  
 PRIOR APPLICATION NUMBER: 60/189,691  
 PRIOR FILING DATE: 2000-03-15  
 PRIOR APPLICATION NUMBER: 09/277,716  
 PRIOR FILING DATE: 1999-03-26  
 PRIOR APPLICATION NUMBER: 08/757,046  
 PRIOR FILING DATE: 1996-11-25  
 PRIOR APPLICATION NUMBER: 08/597,274  
 PRIOR FILING DATE: 1996-02-06  
 PRIOR APPLICATION NUMBER: 08/908,909  
 PRIOR FILING DATE: 1997-08-08  
 PRIOR APPLICATION NUMBER: 08/990,103  
 PRIOR FILING DATE: 1997-12-12  
 NUMBER OF SEQ ID NOS: 33  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 32  
 LENGTH: 238  
 TYPE: PRF  
 ORGANISM: *Ptilosarcus gurneyi*  
 US-11-175-766-32

Query Match 39.0%; Score 473.5; DB 7; Length 238;  
 Best Local Similarity 45.5%; Pred. No. 8.3e-39;  
 Matches 97; Conservative 40; Mismatches 73; Indels 3; Gaps 3;

5 KNV-IKEFMRFKVMEGTVNGHEFEIEGEGRPRYEGHNTVKLKTKGGLPFAMDILSP 63  
 7 KNTGLKEIMSAKASVSGIYNNHVFMEGFGKGNVLFGNQLMQRVTYKGGPLPFADIVSI 66  
 64 OFQYSGKYVVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTODSSLQDGCFTIYKYK 123  
 67 AFQYGNRTFTKYDDIADYFVQSPFAGFYERNLRFEDGALVDIRSDISLEDKCFHYKYE 126  
 124 FIGVNPSPDGPVWQKKTGMGEASTELYPKDGVLKGEIHKALKLKDGGHYLVEFKSIYMA 183  
 127 YRNGPSPNGPVWQKAILGMEPSFEVYVWNSGVLVGEVDLVYKLSGNTYSCHMKTFYRS 186  
 184 KKPV-QLPGYVYVDSKLDITSHNEDYTIYEQYE 215  
 187 KGGVKEPPEYHFIHRLKLT-YVEBGSFVEQHE 218

RESULT 12  
 US-11-179-411-27  
 Sequence 27, Application US/11179411  
 Publication No. US20050266491A1

GENERAL INFORMATION:  
 APPLICANT: Bryan, Bruce  
 APPLICANT: Szent-Gyorgyi, Christopher  
 APPLICANT: Szczeniak, William  
 TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING TI  
 TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
 FILE REFERENCE: 24729-128  
 CURRENT APPLICATION NUMBER: US/11/179,411  
 CURRENT FILING DATE: 2005-07-12  
 PRIOR APPLICATION NUMBER: US/09/808,898  
 PRIOR FILING DATE: 2001-03-15  
 PRIOR APPLICATION NUMBER: 60/189,691  
 PRIOR FILING DATE: 2000-03-15  
 PRIOR APPLICATION NUMBER: 09/277,716  
 PRIOR FILING DATE: 1999-03-26  
 PRIOR APPLICATION NUMBER: 08/757,046  
 PRIOR FILING DATE: 1996-11-25  
 PRIOR APPLICATION NUMBER: 08/597,274  
 PRIOR FILING DATE: 1996-02-06  
 PRIOR APPLICATION NUMBER: 08/908,909  
 PRIOR FILING DATE: 1997-08-08  
 PRIOR APPLICATION NUMBER: 08/990,103  
 PRIOR FILING DATE: 1997-12-12  
 NUMBER OF SEQ ID NOS: 33  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 27  
 LENGTH: 233  
 TYPE: PRF  
 ORGANISM: *Renilla reniformis*  
 US-11-179-411-27

Query Match 34.8%; Score 422.5; DB 7; Length 233;  
 Best Local Similarity 40.5%; Pred. No. 6.9e-34;  
 Matches 90; Conservative 38; Mismatches 93; Indels 1; Gaps 1;

1 MRSSKNVYKEFMRFKVMEGTVNGHEFEIEGEGRPRYEGHNTVKLKTKGGLPFAMDILSP 60  
 1 MDLAKLGLKEVMPKINLEGLVGDHAFSMEGVGEGNILEGTQVXKISVTYGAPLPFAFDI 60  
 61 LSPQYSGKYVVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTODSSLQDGCFTIY 120  
 61 VSAVAFSYGNRAVYGYPEISDYFLQSPBSFTYERINIRYDGGTAIYKSDISLEDKCFIY 120  
 121 KYKFIQVNPSPDGPVWQKKTGMGEASTELYPKDGVLKGEIHKALKLKDGGHYLVEFKSI 180  
 121 NDFPKADLARMGPVWQDILVGMQPSIESYMTVNTVTSIGECIIAFKLTQTKGHFYHMRV 180



/ PRIOR FILING DATE: 1997-12-12  
 / NUMBER OF SEQ ID NOS: 33  
 / SOFTWARE: FastSeq for Windows Version 4.0  
 / SEQ ID NO 33  
 / LENGTH: 233  
 / TYPE: PRT  
 / ORGANISM: Renilla reniformis murein  
 US-11-175-766-33

Query Match 33.8%; Score 410.5; DB 7; Length 233;  
 Best Local Similarity 40.1%; Pred. No. 9.9e-33;  
 Matches 89; Conservative 37; Mismatches 95; Indels 1; Gaps 1;

QY	1	MRSKNVKEFMKFKVMGEGTVNGHEFEIEGEGGRRPYEGHNTVKLKVTKGSPFPAMD	60
DB	1	MDLAKLGLKEVMPTKINLEGLVGDHAFSMEGVGEGNILEGTQEVKISVTKGAPLPAPFD	60
QY	61	LSPOFGYGSKVYKHPADIPDYKLSFPPGFKWERVMNFEDEGVVTQTDDSLQDGCFTY	120
DB	61	VSVAFSYGNRAVYGYPEEISDYFLQSPFPGFTYERNIRYQDGGTAIVDSDISLEDGKFLV	120
QY	121	KVKEIGVNFPSDGPVWQKTMGWEASTERLYPRDGVTKGEIHKALKDKGHHYLVFEKSI	180
DB	121	NVDFKADDLRDMGSPVMQDVGMPSYESMTNTVTSVIGECIIAFKLDQTKDFTYHRTV	180
QY	181	YNAKKPVQ-LPGYVVYDSKLDITSHNEDYTIIVEQYERTEGRH	221
DB	181	YKSKKPEVTMLPHFIQHDLVKTNVDTASGVVQHEPALAAH	222

Search completed: January 11, 2006, 02:06:56  
 Job time : 8 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 12, 2006, 09:05:37 ; Search time 3691 Seconds  
(without alignments)  
3465.125 Million cell updates/sec

Title: US-10-006-922a-12  
Sequence: 1214  
1 MRSSKNVKEFRFRFVMEG.....EDYIVQEYERTEGRHDFL 225

## Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODE=frame\_plus.model -DBV=rlp  
-Q/cgnt\_1/USPTO.spool\_p/US10006922/runat\_10012006\_162408\_18921/app\_query\_faeta\_1.391  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -KOOCL=0 -KOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdt -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=prc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922.ccgnt\_1.1.7415.ccgntat\_10012006\_162408\_18921 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database:

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_in: \*  
3: gb\_env: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pr: \*  
9: gb\_ro: \*  
10: gb\_stb: \*  
11: gb\_sy: \*  
12: gb\_un: \*  
13: gb\_vl: \*  
14: gb\_hcg: \*  
15: gb\_pl: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	678	6	AR183915 Sequence
2	1214	100.0	678	6	AX172854 Sequence
3	1214	100.0	678	6	AX207715 Sequence

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	FEATURES	ORIGIN	ALIGNMENT Scores:	Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
AR183915	AR183915	Sequence 6 from patent US 6342379.	AR183915	AR183915.1	GI:20227884	Unknown.	Unclassified.	1 (bases 1 to 678)	Telen, R.Y. and Gonzalez, J.E. III.	Detection of transmembrane potentials by optical methods	Patent: US 6342379-A 6 29-JAN-2002;	1. 678 /organism="unknown" /mol_type="unassigned DNA"	Alignment Scores: Pred. No.: 1,4e-114 Score: 1214.00 Percent Similarity: 100.00% Best Local Similarity: 100.00% Query Match: 100.00%	678	225	0	0	0	0	678	225	0	0	0	

AX233581 Sequence  
AX370404 Sequence  
AX824725 Sequence  
AF168419 Discosoma  
AR636082 Sequence  
AR670157 Sequence  
AX463698 Sequence  
AX233584 Sequence  
AX233627 Sequence  
AR527331 Sequence  
CO882115 Sequence  
CO882117 Sequence  
AX463702 Sequence  
AX823860 Sequence  
AX569779 Cloning v  
AX663075 Sequence  
AX370406 Sequence  
AX348043 Sequence  
AX353910 Sequence  
AF679107 Discosoma  
AF545828 Discosoma  
AX370408 Sequence  
AX824732 Sequence  
AX679106 Discosoma  
AX824731 Sequence  
AX666133 Sequence  
CO849509 Sequence  
AY818375 Cloning v  
DQ005468 Cloning v  
AY818373 Cloning v  
AY640628 SiRNA vec  
AY440625 SiRNA vec  
AB51284 Cloning v  
AY613997 Cloning v  
AY440634 SiRNA vec  
AY640633 SiRNA vec  
AY440630 SiRNA vec  
CO849511 Sequence  
CS018253 Sequence  
CS018255 Sequence  
CS018254 Sequence

US-10-006-922a-12 (1-225) x AR183915 (1-678)

QY 1 MecArgSerSerLybAaenValIleLygLuPheMetArgPheLybValArgMetGluGly 20  
DB 1 ATGAGGCTCTCCAAAGAAAGTTATCAAGAGTTCAAGAGTTTAAGGTTTCGATCGAAGAGA 60

QY 21 ThrValaengLybHsGluPheGluIleGluGluGlyGluGluGlyGluGluGlyGluGly 40  
DB 61 ACGGTCATGGGCGACAGATTGAAATAGAGGCGAAGAGGAGGAGGCGCATACGAAGGC 120

QY 41 HisAenThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrrAspIle 60  
DB 121 CACAAATACCGTAAGACTTAAGGTAAACCAAGGGGAGCCTTTGCCATTTGCTTGGGATTT 180

QY 61 LeuSerProGluInPheGluInTyrgLySerLybValTyValLybHisProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTTCAATATGGAAGCAAGGTATATGTCAAGCACCTTGGCCGACATACCA 240

QY 81 AspTyrlLybLybLeuSerPheProGluGlyPheLybSTrPGluArgValMetAsnPheGlu 100  
DB 241 GACATTAATAAAGCTGTCAATTCCTTAAGGATTTAATGGAAAGGAGTCAATGAACCTTTGAA 300

QY 101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyrr 120  
DB 301 GACGGTGGCGTCTGTAAGTAAACCAAGATTCGATTTGCAAGAGTGGCTGTTTCACTTAC 360

QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGluInLybThr 140  
DB 361 AAGGTCAAAGTTCATTTGGCGTGAATCTTCCGATGAGACCTGTTATCAAAAGAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLybGlyGlu 160  
DB 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCTGATAGGCGTGTGAAAGAGAG 480

QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrrLeuValGluPheLybSerIle 180  
DB 481 ATTCAATAAGGCTCTGAAGCTGAAGACCGTGTGATTAACCTAAGTTCAAAAGATTT 540

QY 181 TyrrMetAlaLybLybProValGluLeuProGlyTyrrTyrrTyrrValAspSerLybLeuAsp 200  
DB 541 TACATGGCAAGAGACCTGTGACGCTACAGGCTCTACTATGTTGACTCCAAACTGGAT 600

QY 201 IleThrSerHisAenGluAspTyrrThrIleValGluGluInTyrrGluArgThrGluGlyArg 220  
DB 601 ATTAACAAGCCACAAGAACTATACATCGTTAGCGATATGAAGAACGAGGAGAGCC 660

QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCTT 675

RESULT 2  
AX172854 678 bp DNA linear PAT 03-JUN-2001  
LOCUS Sequence 6 from Patent WO0142211.  
DEFINITION AX172854  
ACCESSION AX172854  
VERSION AX172854.1 GI:14597903  
KEYWORDS  
SOURCE  
ORGANISM  
Discosoma sp.  
Discosoma sp.  
Bukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
REFERENCE  
AUTHORS Tsien, R.Y. and Gonzalez, J.E.  
TITLE Detection of transmembrane potentials by optical methods  
JOURNAL Patent: WO 0142211-A 6 14-JUN-2001;  
The Regents of the University of California (US)  
FEATURES  
location/Qualifiers  
1..678  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"

ORIGIN

Alignment Scores:  
Pred. No.: 1.4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AX172854 (1-678)

QY 1 MecArgSerSerLybAaenValIleLygLuPheMetArgPheLybValArgMetGluGly 20  
DB 1 ATGAGGCTCTCCAAAGAAAGTTATCAAGAGTTCAAGAGTTTAAGGTTTCGATCGAAGAGA 60

QY 21 ThrValaengLybHsGluPheGluIleGluGluGlyGluGluGlyGluGluGlyGluGly 40  
DB 61 ACGGTCATGGGCGACAGATTGAAATAGAGGCGAAGAGGAGGAGGCGCATACGAAGGC 120

QY 41 HisAenThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrrAspIle 60  
DB 121 CACAAATACCGTAAGACTTAAGGTAAACCAAGGGGAGCCTTTGCCATTTGCTTGGGATTT 180

QY 61 LeuSerProGluInPheGluInTyrgLySerLybValTyValLybHisProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTTCAATATGGAAGCAAGGTATATGTCAAGCACCTTGGCCGACATACCA 240

QY 81 AspTyrlLybLybLeuSerPheProGluGlyPheLybSTrPGluArgValMetAsnPheGlu 100  
DB 241 GACATTAATAAAGCTGTCAATTCCTTAAGGATTTAATGGAAAGGAGTCAATGAACCTTTGAA 300

QY 101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyrr 120  
DB 301 GACGGTGGCGTCTGTAAGTAAACCAAGATTCGATTTGCAAGAGTGGCTGTTTCACTTAC 360

QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGluInLybThr 140  
DB 361 AAGGTCAAAGTTCATTTGGCGTGAATCTTCCGATGAGACCTGTTATCAAAAGAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLybGlyGlu 160  
DB 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCTGATAGGCGTGTGAAAGAGAG 480

QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrrLeuValGluPheLybSerIle 180  
DB 481 ATTCAATAAGGCTCTGAAGCTGAAGACCGTGTGATTAACCTAAGTTCAAAAGATTT 540

QY 181 TyrrMetAlaLybLybProValGluLeuProGlyTyrrTyrrTyrrValAspSerLybLeuAsp 200  
DB 541 TACATGGCAAGAGACCTGTGACGCTACAGGCTACTATGTTGACTCCAAACTGGAT 600

QY 201 IleThrSerHisAenGluAspTyrrThrIleValGluGluInTyrrGluArgThrGluGlyArg 220  
DB 601 ATTAACAAGCCACAAGAACTATACATCGTTAGCGATATGAAGAACGAGGAGAGCC 660

QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCTT 675

RESULT 3  
AX207715 678 bp DNA linear PAT 31-AUG-2001  
LOCUS Sequence 13 from Patent WO0157242.  
DEFINITION AX207715  
ACCESSION AX207715  
VERSION AX207715.1 GI:1542399  
KEYWORDS  
SOURCE  
ORGANISM  
Discosoma sp.  
Discosoma sp.  
Bukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
REFERENCE  
AUTHORS Stack, J.H., Whitney, M., Cubitt, A.B. and Pollok, B.A.  
TITLE Methods of protein destabilization and uses thereof

JOURNAL Patent: WO 0157242-A 13 09-AUG-2001;  
Aurora Biosciences Corporation (US)  
FEATURES Location/Qualifiers  
source 1..678  
/organism="Discoosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"

ORIGIN

Alignment Scores:  
Pred. No.: 1.4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-006-922a-12 (1-225) x AX207715 (1-678)

QY 1 MetATGSeSerLySeAenValIleLySGIuphMeArGPhelySeValAArgMeGluGly 20  
DB 1 ATGAGGTCTTCCAAAGATTTTATCAAGAGCTTCAAGAGTTTAAGGTTTGAAGGATGGAAGA 60  
QY 21 ThrValaenGlyVHleGluPheGluIleGluGlyGluGlyAArgProTyArgGly 40  
DB 61 ACGGTCAATGGGCAAGATTGAAATAGAAAGCGAAGAGGAGGCGCATACGAAGGC 120  
QY 41 HisAenThValLyLeuLySeValThrLySGIyProLeuProPheAlaTPaPile 60  
DB 121 CACAAATACCGTAAGCTTAAAGGTAAACCAAGGGGAGCCTTGCATTGCTGGGATATT 180  
QY 61 LeuSeProGlnPheGlnTyArgLySeLyValTyValLySHisProAlaAspIlePro 80  
DB 181 TTGTCCACCAATTTTCAGTATGGAACCAAGGTATATGTCAAGCACTCTGCCGACATACCA 240  
QY 81 AspTyLyLeuLySeuSerPheProGluGlyPheLySTrpGluArgValIleAenPheGlu 100  
DB 241 GACTATTAATAAGGTGCATTCTCGTAAGGATTTAAATGGGAAAGGTCATGAACITTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnaAspSerLeuGlnaAspGlyCysPheIleTy 120  
DB 301 GACGGTGGGCGTCTTACTGTAAACCCAGGATTCAGTTTCAGAGATGGCTGTTCACTTAC 360  
QY 121 LySeValLySePheIleGlyValaenPheProSeAspGlyProValMetGlnLySlyThr 140  
DB 361 AAGGTCAAGTCTTATGGCGTAACCTTCTCCGATGGAACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrgluArgLeuTyProArgAspGlyValLeuLySGIyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCGTATGGCGGTGTAAGAGAGAG 480  
QY 161 IleHisLyLeuAlaLeuLySeuLySeuAspGlyGlyVHleTyTrLeuValGluPheLySerIle 180  
DB 481 ATTCATTAAGGCTCTGAAGCTGAAGAGCGTGTCACTTACCTAGTTGAATTCAAAGATATT 540  
QY 181 TyrMetAlaLySeLySeProValGlnLeuProGlyTyTrTyTrValaAspSerLySeuAsp 200  
DB 541 TACATGGCAAAAGAGCTGTGACGCTACCAAGGTACTACATGTGTGACTCAAAAGTGAT 600  
QY 201 IleThSerHisaenGluAspTyTrThrIleValGluGlnTyTrGluArgThrgluGlyArg 220  
DB 601 ATTAACAAGCCACAAGAGACTATACATCGTTGAGAGATGAAAGAACCGAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTCTT 675

RESULT 4  
AX233581 678 bp DNA linear PAT 11-SEP-2001  
LOCUS AX233581  
DEFINITION Sequence 5 from Patent W00162919.  
ACCESSION AX233581  
VERSION AX233581.1 GI:15593305

KEYWORDS  
SOURCE Discoosoma sp.  
ORGANISM Discoosoma sp.  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discoosoma.

REFERENCE  
AUTHORS Nelson, D., Zamatra, E. and Tsien, R.  
TITLE Modified fluorescent proteins  
JOURNAL Patent: WO 0162919-A 5 30-AUG-2001;  
Aurora Biosciences Corporation (US)  
FEATURES Location/Qualifiers  
source 1..678  
/organism="Discoosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"  
/note="red"

ORIGIN

Alignment Scores:  
Pred. No.: 1.4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-006-922a-12 (1-225) x AX233581 (1-678)

QY 1 MetATGSeSerLySeAenValIleLySGIuphMeArGPhelySeValAArgMeGluGly 20  
DB 1 ATGAGGTCTTCCAAAGATTTTATCAAGAGCTTCAAGAGTTTAAGGTTTGAAGGATGGAAGA 60  
QY 21 ThrValaenGlyVHleGluPheGluIleGluGlyGluGlyAArgProTyArgGly 40  
DB 61 ACGGTCAATGGGCAAGATTGAAATAGAAAGCGAAGAGGAGGCGCATACGAAGGC 120  
QY 41 HisAenThValLyLeuLySeValThrLySGIyProLeuProPheAlaTPaPile 60  
DB 121 CACAAATACCGTAAGCTTAAAGGTAAACCAAGGGGAGCCTTGCATTGCTGGGATATT 180  
QY 61 LeuSeProGlnPheGlnTyArgLySeLyValTyValLySHisProAlaAspIlePro 80  
DB 181 TTGTCCACCAATTTTCAGTATGGAACCAAGGTATATGTCAAGCACTCTGCCGACATACCA 240  
QY 81 AspTyLyLeuLySeuSerPheProGluGlyPheLySTrpGluArgValIleAenPheGlu 100  
DB 241 GACTATTAATAAGGTGCATTCTCGTAAGGATTTAAATGGGAAAGGTCATGAACITTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnaAspSerLeuGlnaAspGlyCysPheIleTy 120  
DB 301 GACGGTGGGCGTCTTACTGTAAACCCAGGATTCAGTTTCAGAGATGGCTGTTCACTTAC 360  
QY 121 LySeValLySePheIleGlyValaenPheProSeAspGlyProValMetGlnLySlyThr 140  
DB 361 AAGGTCAAGTCTTATGGCGTAACCTTCTCCGATGGAACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrgluArgLeuTyProArgAspGlyValLeuLySGIyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCGTATGGCGGTGTAAGAGAGAG 480  
QY 161 IleHisLyLeuAlaLeuLySeuLySeuAspGlyGlyVHleTyTrLeuValGluPheLySerIle 180  
DB 481 ATTCATTAAGGCTCTGAAGCTGAAGAGCGTGTCACTTACCTAGTTGAATTCAAAGATATT 540  
QY 181 TyrMetAlaLySeLySeProValGlnLeuProGlyTyTrTyTrValaAspSerLySeuAsp 200  
DB 541 TACATGGCAAAAGAGCTGTGACGCTACCAAGGTACTACATGTGTGACTCAAAAGTGAT 600  
QY 201 IleThSerHisaenGluAspTyTrThrIleValGluGlnTyTrGluArgThrgluGlyArg 220  
DB 601 ATTAACAAGCCACAAGAGACTATACATCGTTGAGAGATGAAAGAACCGAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225

Db 661 CACCATCTGTCCTT 675

RESULT 5  
LOCUS AX370404 678 bp DNA linear PAT 16-FEB-2002  
DEFINITION Sequence 1 from Patent WO0196373.  
ACCESSION AX370404  
VERSION AX370404.1 GI:18857490  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.

REFERENCE  
1 Pradkov, A. F. and Tereshikh, A.  
Autors  
TITLE Fluorescent timer proteins and methods for their use  
JOURNAL Patent: WO 0196373-A 1 20-DEC-2001;  
Clontech Laboratories Inc. (US)  
FEATURES  
SOURCE  
1. 678  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="variant of sequence from *Discosoma* sp."

ORIGIN

Alignment Scores:  
Pred. No.: 1.4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AX370404 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGGCTCTCTCCAGAGAGCTCATCAAGAGATTCAAGCTTCAAGTCCGATGAGGGC 60  
Qy 21 ThrValAsnGlyYHieGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGGAACGGCCACGAGATCGAGGCGAGGCGAGGCGCGCCCTACGAGGGC 120  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTTPAspIle 60  
Db 121 CACAAACCGGAACTGAAGGTGACCAAGGCGGCGCCCTGCTGCTGGGACATC 180  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCACTTCCATACGGCTCCCAAGGTGATCGAAGCACCCCGCGACATCCCC 240  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTTPGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAATTCGAG 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGCGCGCGTGGTGAACCGTGAACCTCTCTGCAAGAGCGGTGCTTCACTTAC 360  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTGAAGTTCACTGGCGCTGAATCTTCCCAAGCGCCCGGTGATGACAAAGAGACC 420  
Qy 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLysLysGlyGlu 160  
Db 421 ATGGGCTGGAGGCTCTCAACGAGCGCTGTACCCCGCAAGGCGTCTGTAAGGGCAG 480  
Qy 161 IleHisLysValAlaLeuLysLeuLysAspGlyGlyHisTyrTrieuValGluPheLysSerIle 180  
Db 481 ATCCCAAGGCGCTGAAAGCTGAAGGACGCGCGCTGCTGCTGAGTTCGAAGTCCATC 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

Db 541 TACATGGCAAGAGCCCGTGCACCTGCCGCTACTACTACGTGACTCCAAAGTGGAC 600

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACTCTCCCAACAGAGACTACACATCGTGAAGAGTACGAGCGCACCGAGGCGCGC 660

Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCACTGTCTCTG 675

RESULT 6  
LOCUS AX824725 678 bp DNA linear PAT 11-DEC-2003  
DEFINITION Sequence 7 from Patent WO02068459.  
ACCESSION AX824725  
VERSION AX824725.1 GI:39750591  
KEYWORDS  
SOURCE  
ORGANISM  
Discosoma sp.  
Discosoma sp.  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.

REFERENCE  
1 Non aggregating fluorescent proteins and methods for using the same  
Autors  
TITLE  
JOURNAL Patent: WO 02068459-A 7 06-SEP-2002;  
FEATURES  
SOURCE  
1. 678  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"

ORIGIN

Alignment Scores:  
Pred. No.: 1.4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AX824725 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGGCTCTCTCCAGAGAGCTCATCAAGAGATTCAAGCTTCAAGTCCGATGAGGGC 60  
Qy 21 ThrValAsnGlyYHieGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGGAACGGCCACGAGATCGAGGCGAGGCGAGGCGCGCCCTACGAGGGC 120  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTTPAspIle 60  
Db 121 CACAAACCGGAACTGAAGGTGACCAAGGCGGCGCCCTGCTGCTGGGACATC 180  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCACTTCCATACGGCTCCCAAGGTGATCGAAGCACCCCGCGACATCCCC 240  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTTPGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAATTCGAG 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGCGCGCGTGGTGAACCGTGAACCTCTCTGCAAGAGCGGTGCTTCACTTAC 360  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTGAAGTTCACTGGCGCTGAATCTTCCCAAGCGCCCGGTGATGACAAAGAGACC 420  
Qy 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLysLysGlyGlu 160



Df	421	ATGGGCTGGAGGCCCTTCCACCGAGCCCTTACCCTCCCGCAGCGGCGTGCTGAAGGGCGAG	480
Oy	161	11glttalyalaleulybleulybaapgllyglylhlsttyrleuva1glnphelyssertlle	180
Df	481	ATCCACAAGGCCCTGMAAGCTGAAGAGACGGGGGCACACTACTGTGGAGTTCAAGTCATC	540
Oy	181	TyrMetAlaIaYelYelSPoVa1glnLeuProGLYTYTYTYTYTYValaIaPserLYleuaAp	200
Df	541	TACATGGCCAAAGAAGCCCGCTGACAGCTGCCCGCGCATCTACACGTGACCTCCAAGCTGGAC	600
Oy	201	11eHseRxtlaangluAapRYTrThleVa1glnGlnIntYrgluAgTrhglnGlyArg	220
Df	601	ATCACCTTCCACACAGAGACTACACCATCTGTGGAGACGTRCAGACGACACCGAGGCCGC	660
Oy	221	H1eHstleupheleu	225
Df	661	CACCACTGTTCTCG	675
RESULT 7			
LOCUS	AF168419	859 bp mRNA linear INV 27-JUL-2001	
DEFINITION	Discosoma sp. fluorescent protein FP583 mRNA, complete cds.		
ACCESSION	AF168419		
VERSION	AF168419.2	GI:7105733	
KEYWORDS			
SOURCE	Discosoma sp.		
ORGANISM	Discosoma sp.		
REFERENCE	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomatidae; Discosoma. 1 (bases 1 to 859)		
AUTHORS	Matz,M.V., Pradkov,A.F., Labas,Y.A., Savitsky,A.P., Zaraisky,A.G., Markelov,M.L. and Lukyanov,S.A. Fluorescent proteins from nonbioluminescent Anthozoa species Nat. Biotechnol. 17 (10), 969-973 (1999) 10504696		
JOURNAL	2 (bases 1 to 859)		
PUBMED	Matz,M.V., Pradkov,A.F., Labas,Y.A., Savitsky,A.P., Zaraisky,A.G., Markelov,M.L. and Lukyanov,S.A.		
AUTHORS	Direct Submission Submitted (14-JUL-1999) Institute of Bioorganic Chemistry RAS, Mklukho-Maklaya 16/10, Moscow 117871, Russia 3 (bases 1 to 859)		
JOURNAL	Matz,M.V., Pradkov,A.F., Labas,Y.A., Savitsky,A.P., Zaraisky,A.G., Markelov,M.L. and Lukyanov,S.A. Direct Submission Submitted (25-FEB-2000) Institute of Bioorganic Chemistry RAS, Mklukho-Maklaya 16/10, Moscow 117871, Russia Sequence update by submitter On Feb 25, 2000 this sequence version replaced gi:6090866. Location/Qualifiers		
REMARK	COMMENT		
FEATURES	source		
	1..859	/organism="Discosoma sp."	
		/mol_type="mRNA"	
		/db_xref="taxon:86600"	
	54..731	/note="GFP-like protein; orange-red"	
		/codon_start=1	
		/product="fluorescent protein FP583"	
		/protein_id="AAF03369.1"	
		/db_xref="gi:6090867"	
		/translation="MRSSKNV1KEFRKPKVMEGTVNGHEFEIGEGGRPYEGHTNV KLVTQGSPLPPAMDI1LSPFOYGSKVVVKHPADI PDYKKLSFPGFGFMERVMNPFEDGG GLVTVTGDSSLDDGCFIVKVFIGNVPSPSDPVWAKMTMGWEASTERLYPRPDGV1KGSGE IHKLKIKDGGHYLVFERPSIMWAKKPVLDPGYVVDRLDTITSHNEIDYIVEQEYETERT GRHLFL"	
ORIGIN			
Alignment Scores:			
Pred. No.:	1,87e-114	Length:	859
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatch:	0
Query Match:	100.00%	Indels:	0

DB:	2	Gaps:	0
US-10-006-922A-12 (1-225) x AF168419 (1-859)			
QY	1	MetArSerSerIysAsnValIleIysGluPheMetArPheIysValArgMetGluGly	20
Db	54	ATGAGGCTTCCCAAGAAATGTTATCAAGAGAGTTTATGATGAGTTTAAAGTTCGATGGAAGGA	113
QY	21	ThrValIAsnGlyHisGluPheGluIleGluGlyGluGlyValGluGlyArgProTyrGluGly	40
Db	114	ACCGTCATGGGACGAGGTTGAAATGAAAGGGCGAAAGAGAGGGAGGCCATACGAAGGC	173
QY	41	HisAsnThrValIysLeuIysValIleThrIysGlyGlyProLeuProPheIleATPAspIle	60
Db	174	CACAAATACCTTAAGCTTAAAGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGAAATT	233
QY	61	LeuSerProGluPheGluIleTyrGlySerIysValIleTyrValIysHisProAlaAspIlePro	80
Db	234	TTGTCAACCAATTTTCAATGATGGAAGCAAGATATATGTAACGACACCTGCCGATACCA	293
QY	81	AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu	100
Db	294	GACATATATAAGCTGCTCATTTCCGAAAGATTAAATGGAAAGGGTCATGAACCTTGAA	353
QY	101	AspGlyGlyValValIleThrValIleThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	354	GACGCTGACGCTGATTCTGTAAACCAAGATTTCCAGTTTGCAGATGGCTGTTTCATCTAC	413
QY	121	LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGluIysIysThr	140
Db	414	AAAGTCAAGTTTCATTTGGCGTGAACCTTCTTCGATGACCTGTATGCAAAAGAAAGACA	473
QY	141	MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValIleIysGlyGlu	160
Db	474	ATGGGCTGGGAAGCACACATGAGCCGTTGATCTCTGTATGGCGGTTGAAAGAGAG	533
QY	161	IleHisIysIleLeuIysLeuIysAspGlyGlyHisTyrIleValGluPheIysSerIle	180
Db	534	ATTATATAAGCTCTGAAGCTGAAAGACGGTGGTCAATTAACCTAGTTGAATCAAAAGTATT	593
QY	181	TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp	200
Db	594	TACATGGCAAGAAAGCCCTGGACACTACAGGGTACTACTATGTTGACTCCAAACTGGAT	653
QY	201	IleThrSerHisAsnGluAspTyrThrIleValIleGluGlnTyrGluArgThrGluGlyArg	220
Db	654	ATAACAAGCCCAACAAGAACTATACATCGTTAGACGATATGAAAGAACCGAGGACGC	713
QY	221	HisHisLeuPheLeu 225	
Db	714	CACCATCTGTTCTCT 728	
RESULT 8			
AR636082		859 bp	DNA linear PAT 14-FEB-2005
LOCUS	AR636082		
DEFINITION	Sequence 11 from patent US 6852849.		
ACCESSION	AR636082		
VERSION	AR636082.1		
KEYWORDS			
SOURCE	.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 859)		
TITLE	Tsien, R.Y. and Campbell, R.E.		
JOURNAL	Non-oligomerizing tandem fluorescent proteins		
FEATURES	Patent: US 6852849-A 11 08-FEB-2005;		
source	The Regents of the University of California; Oakland, CA		
	Location/Qualifiers		
	1..859		
ORIGIN	/organism="unknown"		
	/mol_type="genomic DNA"		

Alignment Scores:		Pred. No.: 1,87e-114	Length: 859
Score:	1214.00	Matches: 225	
Percent Similarity:	100.00%	Conservative: 0	
Best Local Similarity:	100.00%	Mismatches: 0	
Query Match:	100.00%	Indels: 0	
DB:	6	Gaps: 0	
US-10-006-922a-12 (1-225) x AR636082 (1-859)			
QY	1	MeCARGSerSerLySaSnVAlIleLySgIuPheMeCARGPheLyVaIArgMeCgIuGly	20
DB	54	ATGAGTCTTCCAAAGAAATGTTATCAAGGAGTTCAATGAGGTTTAAGGTTCCGATGGAAGCA	113
QY	21	ThrValaSnGlyHhIeGluPheGluIleGluGlyGluGlyGluGlyYArgProTyTrGluGly	40
DB	114	ACGGCAATGGGCAAGATTGAAATAGAAAGCAAGAGAGGAGGCAATACCAAGGC	173
QY	41	HhSaSnThrValLyLeuLyVaIThrLySgIyGlyProLeuProPheAlaTrpAspIle	60
DB	174	CACAAATACCGTAAACCTTAAGGTAAACCAAGGGGGAACCTTCCATTGCTTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTyTrGlySerLyVaIlyrValLyHhIeProAlaAspIlePro	80
DB	234	TTGTCAACCAATTTCAAGTATGAAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA	293
QY	81	AspTyTrLyLyLeuSerPheProGluGlyPheLySTrPGluArgVaIMetAsnPhGlu	100
DB	294	GACTATAAAGAGCTGTCAATTTCTTAAGGATTTAAATGGAAAGGCTCATGAACCTTTGAA	353
QY	101	AspGlyGlyValValIThrValIThrGlnAspSerSerLeuGlnAspGlyCyAspHeIleTy	120
DB	354	GACGCTGGCGTGTACTGTAACCAAGATTTCCATTGGCAGAGATGGCTGTTTCACTAC	413
QY	121	LyVaIlySphelIeGlyValaSnPheProSerAspGlyProValMetGlnLyLySThr	140
DB	414	AAGGCAAGTTCATTTGGCGTGAACCTTCTCCATGAGACCTGTATCAAAAGAACACA	473
QY	141	MetGlyTrPGluIaSerThrGluArgLeuTyTrProArgAspGlyValLeuLyGlyIu	160
DB	474	ATGGCTGGGAAGCAGACCTGAGCGTTGTATCTCGTGAAGGCGGTGTGAAGAGAG	533
QY	161	IleHhIySaIaLeuLyLeuLyAspGlyGlyHhIeTyTrLeuValGluPheLySerIle	180
DB	534	ATTGATAAGGCTCTGAAGCTGAAGACGGTGTCAATTACCTAGTGAATTCAAAAGTATT	593
QY	181	TyrMeCAlaLyLySphProValGlnLeuProGlyTyTrTyTrValaAspSerLyLeuAsp	200
DB	594	TACATGGCAAGAACCTGTGCACTACCAAGGTTCTACTATGTGACTCCAAACTGGAT	653
QY	201	IleThrSerHhSaSnGluAspTyTrThrIleValGluGlnTyTrGluArgThhGluGlyArg	220
DB	654	ATTAACAGCCACAAAGAACTATACAAATCGTTGAGAGTATGAAGAACCAAGGAGAGC	713
QY	221	HhIhIeLeuPheLeu 225	
DB	714	CACCATCTGTTCCCTT 728	
RESULT 9			
LOCUS	AR670157	859 bp	DNA linear PAT 13-JUN-2005
DEFINITION	Sequence 11 from patent US 6900304.		
ACCESSION	AR670157		
VERSION	AR670157.1 GI:67608896		
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 859)		
AUTHORS	Tsien,R.Y., Ting,A.Y. and Zhang,J		
TITLE	Emulsion radiometric indicators of phosphorylation		
JOURNAL	Patent: US 6900304-A 11 31-MAY-2005;		
	The Regents of the University of California; Oakland, CA		

FEATURES		Location/Qualifiers	
SOURCE		1..859	
ORIGIN		/organism="unknown"	
		/mol_type="genomic DNA"	
Alignment Scores:			
Pred. No.:	1,87e-114	Length:	859
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x AR670157 (1-859)			
QY	1	MeCARGSerSerLySaSnVAlIleLySgIuPheMeCARGPheLyVaIArgMeCgIuGly	20
DB	54	ATGAGGCTCTTCCAAAGAAATGTTATCAAGGAGTTCAATGAGGTTTAAGGTTCCGATGGAAGCA	113
QY	21	ThrValaSnGlyHhIeGluPheGluIleGluGlyGluGlyYArgProTyTrGluGly	40
DB	114	ACGGCAATGGGCAAGATTGAAATAGAAAGCAAGAGAGGAGGCAATACCAAGGC	173
QY	41	HhSaSnThrValLyLeuLyVaIThrLySgIyGlyProLeuProPheAlaTrpAspIle	60
DB	174	CACAAATACCGTAAACCTTAAGGTAAACCAAGGGGGAACCTTGGCATTGCTTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTyTrGlySerLySaIlyrValLyHhIeProAlaAspIlePro	80
DB	234	TTGTCAACCAATTTCAAGTATGAAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA	293
QY	81	AspTyTrLyLyLeuSerPheProGluGlyPheLySTrPGluArgVaIMetAsnPhGlu	100
DB	294	GACTATAAAGAGCTGTCAATTTCTTAAGGATTTAAATGGAAAGGCTCATGAACCTTTGAA	353
QY	101	AspGlyGlyValValIThrValIThrGlnAspSerSerLeuGlnAspGlyCyAspHeIleTy	120
DB	354	GACGCTGGCGTGTACTGTAACCAAGATTTCCATTGGCAGAGATGGCTGTTTCACTAC	413
QY	121	LyVaIlySphelIeGlyValaSnPheProSerAspGlyProValMetGlnLyLySThr	140
DB	414	AAGGCAAGTTCATTTGGCGTGAACCTTCTCCATGAGACCTGTATCAAAAGAACACA	473
QY	141	MetGlyTrPGluIaSerThrGluArgLeuTyTrProArgAspGlyValLeuLyGlyIu	160
DB	474	ATGGCTGGGAAGCAGACCTGAGCGTTGTATCTCGTGAAGGCGGTGTGAAGAGAGAG	533
QY	161	IleHhIySaIaLeuLyLeuLyAspGlyGlyHhIeTyTrLeuValGluPheLySerIle	180
DB	534	ATTGATAAGGCTCTGAAGCTGAAGACGGTGTCAATTACCTAGTGAATTCAAAAGTATT	593
QY	181	TyrMeCAlaLyLySphProValGlnLeuProGlyTyTrTyTrValaAspSerLyLeuAsp	200
DB	594	TACATGGCAAGAACCTGTGCACTACCAAGGTTCTACTATGTGACTCCAAACTGGAT	653
QY	201	IleThrSerHhSaSnGluAspTyTrThrIleValGluGlnTyTrGluArgThhGluGlyArg	220
DB	654	ATTAACAGCCACAAAGAACTATACAAATCGTTGAGAGTATGAAGAACCAAGGAGAGC	713
QY	221	HhIhIeLeuPheLeu 225	
DB	714	CACCATCTGTTCTT 728	
RESULT 10			
LOCUS	AX463698	859 bp	DNA linear PAT 15-JUL-2002
DEFINITION	Sequence 12 from Patent WO248338.		
ACCESSION	AX463698		
VERSION	AX463698.1 GI:21886457		
KEYWORDS			
SOURCE	Discozoa sp.		
ORGANISM	Discozoa sp.		

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Lichtenberg-Frat. H.  
Yeast strain for testing the geno- and cytotoxicity of complex environmental contamination  
Patent: WO 0248338-A 12 20-JUN-2002;  
Lichtenberg-Frat, Heila (DE)  
Location/Qualifiers  
source  
1..859  
/organism="Discozoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"  
54..731  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD42147.1"  
/db\_xref="GI:21886458"  
/translation="MRSSKNVYKERMRPKVMRGTVNGHEFEIBEGGRPYEGHNTV  
KLKVTGSPPLPFPAMDILSPQFGSKYVYKHPADIPYKQLSPFGFMRVWNPEDG  
GVVTVTDDSLDGGCFIYKVFIVGNVPSDGVPMQKTMGEASTERLYPRDVLKGB  
IHKALKDGGHYLVFPSIYMAKKPVQLPGYVYDSDKLITSHNEDYTIIVEQYTE  
GRHHLFL"

ORIGIN

Alignment Scores:  
Pred. No.: 1..87e-114 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX463698 (1-859)

QY 1 MetAAGSerSerlyabAnVal11elysgIuphemeArxPheIyValArgMetGluGly 20  
DB 54 ATGAGGCTTCCAGAAATGTTATCAAGAGGTTTCATGAGGTTTAAAGTTGCGATGAGGA 113  
QY 21 ThrValAenGlyVhiGluPheGluIleGluGlyGluGlyGluGlyVarpProTyrgGluGly 40  
DB 114 AGCGTCAATGGGCAACAGTTTGAAATAGAAAGGAGAGAGGAGGCGCATACGAAGC 173  
QY 41 H1aenThValIyLeuIyValThrlYsgIyGlyProLeuProPhealAtPaaPile 60  
DB 174 CACAAACCGTAAGCTTAAGTTAACCAAGGGGGAAGCTTGGCATTTGGATATC 233  
QY 61 LeuSerProGlnPheGlnTyrgIySerlyValTyrrValIyVhiSerProAlaAapPilePro 80  
DB 234 TTGTCAACCAATTTCAATGATGAGCAAGGTATATGTCAGACACCTGCCGACATACCA 293  
QY 81 AapTyrrLyLeuSerPheProGluGlyPheLyTrpGluIuArgValMetAenPheGlu 100  
DB 294 GACTATAAAGAGTGCATTTCTGAAGATTTAAATGGGAAAGGTCATGAATTTGAA 353  
QY 101 AapGlyGlyValValThrlYsgIySerSerSerleuGlnAapGlyCyPheIleTyrr 120  
DB 354 GACGGGGGGTGGTTCATGTAACCAAGATTCAGATTGGAGATGGCTGTTTCATCTAC 413  
QY 121 LyValIyAenPheIleGlyValAenPheProSerAapGlyProValMetGlnIySlyTrh 140  
DB 414 AAGGTCAAGTTCAATGGCGTGAACCTTCTTCGATGAGACCTGTATGCAAAAGAGAC 473  
QY 141 MetGlyTrpGluIuAserThrgIuArgLeuTyrrProArgAapGlyValIleuLySlyGlu 160  
DB 474 ATGGGCTGGGAAGCCAGACCTGAGCGTTTATCTCTGATAGCGCTGTTGAAAGAGAG 533  
QY 161 H1aIyValIyLeuIySlyLeuIyAapGlyGlyVhiSerTyrrLeuValGluPheLySertile 180  
DB 534 ATTCATTAAGCTCTGAAGCTGAAGAGCGGTGTCATTACTAGTTGAATCAAAAGTAT 593  
QY 181 TyrrMetAlaIyLyPProValGlnLeuProGlyTyrrTyrrTyrrValAapSerlySlyAap 200

DB 594 TACATGGCAAGAAAGCCTGTGACAGCTACAGGTACTACTATGTTGACTCCAAATCGAT 653  
QY 201 I1ethrSerH1aenGluAapTyrrThrlIyValIyGluGlnTyrgIuArgThrgIuGlyArg 220  
DB 654 ATACAAAGCCCAACGAAGACTATACATCGTTGAGCAGTATGAAGAAACGAGGAGCGC 713  
QY 221 H1h1eIuPheLeu 225  
DB 714 CACCATCTGTCTCT 728

RESULT 11  
AX233584  
LOCUS  
DEFINITION  
Sequence 8 from Patent WO0162919.  
ACCESSION  
AX233584  
VERSION  
AX233584.1 GI:15593307  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
other sequences; artificial sequences.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Nelson, D., Zamatra, E. and Tsien, R.  
Modified fluorescent proteins  
Patent: WO 0162919-A 8 30-AUG-2001;  
Aurora Biosciences Corporation (US)  
Location/Qualifiers  
source  
1..681  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Mutant Anthozoan red fluorescent protein"  
1..681  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC69734.1"  
/db\_xref="GI:15593308"  
/translation="MRSSKNVYKERMRPKVMRGTVNGHEFEIBEGGRPYEGHNT  
VLKVTGSPPLPFPAMDILSPQFGSKYVYKHPADIPYKQLSPFGFMRVWNPEDG  
GVVTVTDDSLDGGCFIYKVFIVGNVPSDGVPMQKTMGEASTERLYPRDVLKGB  
IHKALKDGGHYLVFPSIYMAKKPVQLPGYVYDSDKLITSHNEDYTIIVEQYTE  
BSRHLFL"

ORIGIN

Alignment Scores:  
Pred. No.: 3..61e-114 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX233584 (1-681)

QY 1 MetAAGSerSerlyabAnVal11elysgIuphemeArxPheIyValArgMetGluGly 20  
DB 4 GTGAGGAGCAAGCAAGCTGATCAAGAGGTTTCATGAGGTTTAAAGTTGCGATGAGGCG 63  
QY 21 ThrValAenGlyVhiGluPheGluIleGluGlyGluGlyGluGlyVarpProTyrgGluGly 40  
DB 64 ACCGTAAACGCCACAGATTTCAGATCGAGGGGAGGGGAGGGGAGGCCCTTACGAGGGC 123  
QY 41 H1aenThValIyLeuIyValThrlYsgIyGlyProLeuProPhealAtPaaPile 60  
DB 124 CACAAACCGTGAAGCTTAAGGTGACCAAGGGGCGCCCTGCTTCGCTGGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrgIySerlyValTyrrValIyVhiSerProAlaAapPilePro 80  
DB 184 CTGAGCCCCCAGTTCCAGTACGAGCAAGAGGTGTACGTAAAGACACCCCGCCATCATCCC 243  
QY 81 AapTyrrLyLeuSerPheProGluGlyPheLyTrpGluIuArgValMetAenPheGlu 100  
DB 244 GACTACAGAAAGCTGAGCTTCCCGAGGGCTTCAAGTGGGAGAGGCTGATGAATCTCGAG 303

QY 101 AapglYglYvalIValThrValThrglnAapSerSerleuGlnAapGlyCyaphe11eTyr 120  
DB 304 GACGGCGCGGTGTACCGTGAACCAAGACAGACCGCTGCAGAGCGGCTTCATCTAC 363  
QY 121 LysValIysPhe11eGlyValAasnPheProSerAapGlyProValMetGlnIlyserThr 140  
DB 364 AAGGTGAAGTTCATCGGCGTGAACCTCCCAAGCGACGGCCCCGTGATGAGAAAGAAC 423  
QY 141 MetGlyTTPGluAlaSerThrGlnArgLeuTyrProArgAapGlyValIleuIysGly 160  
DB 424 ATGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGGCTGCTGAAGGCGAG 483  
QY 161 ILeHslysaIaleuIlyseuIysAapGlyGlyHsTyrIleuValGlnPheIysSer11e 180  
DB 484 ATCCAAAGGCGCTGAAGCTGAAGAGCGGCGCACTACCTGTGTGAAGTTCAAAGTCCATC 543  
QY 181 TyrMetAlaIlylylyPProValGlnleuProGlyTyrTyrTyrValAapSerIysleuAap 200  
DB 544 TACATGGCCAAAGACCCGTGACGCTGCCGCGCTACTACTAGTGAATCCAAAGCTGAC 603  
QY 201 ILeThrSerHsaAengIuAapTyrThrIleValGlnIlyArgTyrGluArgThrGluIlyArg 220  
DB 604 ATCAACCAAGCCCAACAGAGACTACACATCGTGAGCAGTACAGAGAGACCGAGGCGAG 663  
QY 221 HsHsIleuPheIeu 225  
DB 664 CACCACCTGTTCTTG 678

RESULT 12  
AX233627 713 bp DNA linear PAT 11-SEP-2001  
LOCUS Sequence 51 from Patent WO0162919.  
DEFINITION AX233627  
VERSION AX233627.1 GI:15593330  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences, artificial sequences.

REFERENCE 1  
AUTHORS Nelson,D., Zamaiera,E. and Tsien,R.  
TITLE Modified fluorescent proteins  
JOURNAL Patent: WO 0162919-A 51 30-AUG-2001;  
Aurora Biosciences Corporation (US)  
FEATURES  
SOURCE location/Qualifiers  
1..713  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Mutant Anthozoan red fluorescent protein"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.81e-114 Length: 713  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AX233627 (1-713)

QY 1 MetArgSerSerIysAasnValIleIysGlnPheMetArgPheIysValArgMetGlnIly 20  
DB 22 GTGAGGACACCAAGAGTGTGATCAAGAGTTCATGAGGTGCGCATGAGAGGAG 81  
QY 21 ThrValAengIyHsIleGluPheGluIleGluIlyGluIlyGluIlyArgProTyrGluIly 40  
DB 82 ACCGGAACCGGCGCAAGATTCAGAGTTCAGAGGCGGCGAGGCGAGGCGCTTCACAGAGG 141  
QY 41 HsAasnThrValIysLeuIysValThrIysGlyIlyProLeuProPheAlaTTPAap11e 60  
DB 142 CACAACACCGTGAAGCTTAAGGTGACCAAGGCGGCGCGCTGCTTCGCTGGAGCATTC 201

QY 61 LeuSerProGlnPheGlnTyrGlySerIysValIlyrValIysHsIleProAlaAap11ePro 80  
DB 202 CTGAGCCCCCGATTCAGATAGGAGCAAGGTGTACGTGAAGCAACCCCGGACATCTCC 261  
QY 81 AspTyrIlylyleuSerPheProGluGlyPheIysTTPGluArgValMetAasnPheGlu 100  
DB 262 GACTACAAAGAGCTGAGCTTCCCGAGGCTTCAGAGTGGAGAGAGGTGATGAATCTGAG 321  
QY 101 AapGlyglYvalIValThrValThrglnAapSerSerleuGlnAapGlyCyaphe11eTyr 120  
DB 322 GACGCGCGCGGTGACCGTGAACCAAGACAGACCGCTGCAGACGGCTGCTTCATCTAC 381  
QY 121 LysValIysPhe11eGlyValAasnPheProSerAapGlyProValMetGlnIlyserThr 140  
DB 382 AAGGTGAAGTTCATCGGCGTGAACCTCCCAAGCGACCGCCCCGTGATGACAAAGAAC 441  
QY 141 MetGlyTTPGluAlaSerThrGlnArgLeuTyrProArgAapGlyValIleuIysGly 160  
DB 442 ATGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGGCTGCTGAAGGCGAG 501  
QY 161 ILeHslysaIaleuIlyseuIysAapGlyGlyHsTyrIleuValGlnPheIysSer11e 180  
DB 502 ATCCAAAGGCGCTGAAGCTGAAGAGAGCGGCGCACTACCTGTGTGAAGTTCAAAGTCCATC 561  
QY 181 TyrMetAlaIlylylyPProValGlnleuProGlyTyrTyrTyrValAapSerIysleuAap 200  
DB 562 TACATGGCCAAAGACCCGTGACGCTGCCGCGCTACTACTAGTGAATCCAAAGCTGAC 621  
QY 201 ILeThrSerHsaAengIuAapTyrThrIleValGlnIlyArgTyrGluArgThrGluIlyArg 220  
DB 622 ATCAACCAAGCCCAACAGAGACTACACATCGTGAGCAGTACAGAGAGACCGAGGCGAG 681  
QY 221 HsHsIleuPheIeu 225  
DB 682 CACCACCTGTTCTTG 696

RESULT 13  
AR527331 723 bp DNA linear PAT 08-OCT-2004  
LOCUS AR527331  
DEFINITION Sequence 1 from patent US 6723537.  
ACCESSION AR527331  
VERSION AR527331.1 GI:53914309  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 723)  
AUTHORS Peelle,B.  
TITLE Directed evolution of protein in mammalian cells  
JOURNAL Patent: US 6723537-A 1 20-APR-2004;  
Rigel Pharmaceuticals, Incorporated and Becton, Dickinson and Company; South San Francisco, CA  
FEATURES  
SOURCE location/Qualifiers  
1..723  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.88e-114 Length: 723  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AR527331 (1-723)

QY 1 MetArgSerSerIysAasnValIleIysGlnPheMetArgPheIysValArgMetGlnIly 20  
DB 4 GTGCGTCTCTCCAAAGACGTGATCAAGAGTTCATGCGCTTCAGAGGTGCGCATGAGAGG 63

QY 21 ThrValaAsnglyHlaSglubheglunleagluygluygluyAArgProTyrTgluy 40  
DB 64 ACCGGAAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGGCCGCCCTTACAGAGGC 123  
QY 41 HlaBenthrValaLysleuLysValaThrlYsglygluyProleuProPheAlaTrpAspIle 60  
DB 124 CACAAACACCGTGAAGGTGAAGGTGACCAAGGGCGGCCCTTCCCTCCCTGGAGCATTC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValaTyrValaLysHlaProAlaAspIlePro 80  
DB 184 CTGTCCCCCAAGTTCCAGTACGAGCTTCAGAGGTGTACGTAGACACCCGCCCAACATCCCC 243  
QY 81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTyrGlnArgValaMetAsnPheGlu 100  
DB 244 GACTCAAGAAACGCTGTCTTCCCGAGGGCTTCAAGTGGAGGCGCTGATGAATTCGAG 303  
QY 101 AArgGlyGlyValaValaThrlValaThrglnAspSerSerLeuGlnAspGlyCyAspHeileTyr 120  
DB 304 GACGGCGCGGTGTGAACCGTGAACCCAGAGCTCTCTCCAGAGCGGCTGTGATCATCTAC 363  
QY 121 LysValaLysPheIleGlyValaAsnPheProSerAspGlyProValaMetGlnLysLysThr 140  
DB 364 AAGGTAAAGTTCAATCGGCGGTGAATCTTCCCTCCAGCGGCCCTGTAATGAGAAAGAAC 423  
QY 141 MetGlyTyrGlnLysaSerThrglnArgLeuTyrProArgAspGlyValaLysGlyGlu 160  
DB 424 ATGGGCTGGAGAGCCCTCAACCGAGCGCTGTACCCCGGCGAGCGGCTGTGAAGGGCGAG 483  
QY 161 HlaHlaLysAlaLeuLysLysLeuLysAspGlyGlyHlaTyrLeuValaGlnPheLysSerIle 180  
DB 484 ATCCAAAGAGCCCTGAAGCTGAAGAGAGCGGGCGCCTACTCTGTGTGAAGTTCAAGATATC 543  
QY 181 TyrMetAlaLysLysPheProValaGlnLeuProGlnTyrTyrTyrValaAspSerLysLeuAsp 200  
DB 544 TACATGGCCAAAGAGCCCTGACGCTGCCGCGCTACTACTGAGTCCAGAGCTGAGAC 603  
QY 201 HlaThrSerHlaAsnglyLysPheTyrThrlLeuValaGlnGlnTyrGlnArgThrglnArg 220  
DB 604 ATCACTCCCAACAGAGGAGCTACACCATGTGAGAGAGTACGAGCGCACCGAGGGCGCG 663  
QY 221 HlaHlaLeuPheLeu 225  
DB 664 CACCACCTGTCTCTG 678

RESULT 14  
CQ882115  
LOCUS CQ882115 2721 bp DNA linear PAT 11-OCT-2004  
DEFINITION Sequence 1 from Patent WO2004083445.  
ACCESSION CQ882115  
VERSION CQ882115.1 GI:54034825  
KEYWORDS  
SOURCE .  
ORGANISM synthetic construct  
1  
other sequences; artificial sequences.

REFERENCE  
AUTHORS Chavancy,G., Couble,P., Durand,B., Grenier,A.M., Horard,B.,  
Julien,B., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C.  
TITLE Nucleic acid controlling the expression of a useful polypeptide in  
JOURNAL the posterior silk glands of a lepidoptera and application thereof  
Patent: WO 2004083445-A 1 30-SEP-2004;  
Centre National de la Recherche Scientifique-CNRS (FR); UNIVERSITE  
CLAUDE BERNARD - LYON 1 (FR); Institut National de la Recherche  
Agronomique (INRA) (FR)  
FEATURES  
Source Location/Qualifiers  
1..2721  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Description de la s quence artificielle : s quence  
de fusion"

ORIGIN  
Alignment Scores:

Pred. No.: 1,95e-113 Length: 2721  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatch: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x CQ882115 (1-2721)

QY 1 MetArgSerSerLysAsnValaIleLysGlnPheMetArgPheLysValaArgMetGlnGly 20  
DB 2044 GTGGCTCTCTCCAAAGAGTATCAAGAGGATTCATGCGCTTCAAGGTGCGAGTGAAGGGCC 2103  
QY 21 ThrValaAsnglyHlaSglubheglunleagluygluygluyAArgProTyrTgluy 40  
DB 2104 ACCGTGAACCGGCCACGAGTTGAGATTCAGAGGGCGAGGGCGAGGGCGGCCCTTACAGAGGC 2163  
QY 41 HlaBenthrValaLysleuLysValaThrlYsglygluyProleuProPheAlaTrpAspIle 60  
DB 2164 CACAAACCGTGAAGGTGAAGGTGACCAAGGGCGGCCCTTCCCTTCCCTGGAGCATTC 2223  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValaTyrValaLysHlaProAlaAspIlePro 80  
DB 2224 CTGTCCCCCAAGTTCCAGTACGAGCTTCAGAGGTGTACGTGAAGACACCCGCCGACATCCCC 2283  
QY 81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTyrGlnArgValaMetAsnPheGlu 100  
DB 2284 GACTCAAGAAACGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGGCTGTGAATTCGAG 2343  
QY 101 AArgGlyGlyValaValaThrlValaThrglnAspSerSerLeuGlnAspGlyCyAspHeileTyr 120  
DB 2344 GACGGCGCGGTGTGAACCGTGAACCCAGAGCTCTCTCCAGAGAGCGGCTGTGATCATCTAC 2403  
QY 121 LysValaLysPheIleGlyValaAsnPheProSerAspGlyProValaMetGlnLysLysThr 140  
DB 2404 AAGGTAAATTCATGGCGGTGAATCTTCCCTCCAGCGGCCCTGTAATGAGAAAGAAC 2463  
QY 141 MetGlyTyrGlnLysaSerThrglnArgLeuTyrProArgAspGlyValaLysGlyGlu 160  
DB 2464 ATGGGCTGGAGAGCCCTCAACCGAGCGCTGTACCCCGGAGCGGCTGTGAAGGGCGAG 2523  
QY 161 HlaHlaLysAlaLeuLysLysLeuLysAspGlyGlyHlaTyrLeuValaGlnPheLysSerIle 180  
DB 2524 ATCCAAAGAGCCCTGAAGCTGAAGAGAGCGGCCCTACTCTGTGAGTCCAAAGTCAATC 2583  
QY 181 TyrMetAlaLysLysPheProValaGlnLeuProGlnTyrTyrTyrValaAspSerLysLeuAsp 200  
DB 2584 TACATGGCCAAAGAGCCCTGACGCTGCCGCGCTACTACTGAGTCCAAAGCTGAGAC 2643  
QY 201 HlaThrSerHlaAsnglyLysPheTyrThrlLeuValaGlnGlnTyrGlnArgThrglnArg 220  
DB 2644 ATCACTCCCAACAGAGGAGCTACACCATGTGAGAGAGTACGAGCGCACCGAGGGCGCG 2703  
QY 221 HlaHlaLeuPheLeu 225  
DB 2704 CACCACCTGTCTCTG 2718

RESULT 15  
CQ882117  
LOCUS CQ882117 2772 bp DNA linear PAT 11-OCT-2004  
DEFINITION Sequence 3 from Patent WO2004083445.  
ACCESSION CQ882117  
VERSION CQ882117.1 GI:54034827  
KEYWORDS  
SOURCE .  
ORGANISM synthetic construct  
1  
other sequences; artificial sequences.

REFERENCE  
AUTHORS Chavancy,G., Couble,P., Durand,B., Grenier,A.M., Horard,B.,  
Julien,B., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C.  
TITLE Nucleic acid controlling the expression of a useful polypeptide in  
JOURNAL the posterior silk glands of a lepidoptera and application thereof  
Patent: WO 2004083445-A 3 30-SEP-2004;

Centre National De La Recherche Scientifique-CNRS (FR) ; UNIVERSITE  
CLAUDE BERNARD - LYON 1 (FR) ; Institut National de la Recherche  
Agronomique (INRA) (FR)

FEATRES  
source  
Location/Qualifiers  
1. .2772  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Description de la s quence artificielle : s quence  
de fusion"

ORIGIN

Alignment Scores:

Pred. No.:	2e-113	Length:	2772
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922a-12 (1-225) x CQ882117 (1-2772)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
DB	2095	GTGCGCTCTCCCAAGAACGTATCAAGAGTTCATGCGCTTCAAGTGCGCATGGAGGCG	2154
QY	21	ThrValAngLYHISGLuPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly	40
DB	2155	ACCGGAAACGGCCACGAGTTCAGAGTCGAGGCGGAGGCGGCGCCCTACGAGGCG	2214
QY	41	HISAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60
DB	2215	CACAACACCGTAGAGCTGAAGTGCAGAGGCGGCGCCCTGCGCTTGCGGAGCATC	2274
QY	61	LeuSerProGluPheGluTyrGlySerLysValTyrValLysPheAlaAspIlePro	80
DB	2275	CTGTCCCCCAGTTCACAGTACGGCTCCAGAGGTGATCGTGAAGCACCCCGCACATCCC	2334
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTyrProArgAspGlyValLeuLysGlu	100
DB	2335	GACTACAAAGAGCTGCTCTTCCCGAGGCTTCAAGTGGAGCGCGTATGAACCTCGAG	2394
QY	101	AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyr	120
DB	2395	GACGCGCGCGTGGTACCGTGAACCAAGGATCTCTCCGACGAGCGGCTGCTCATCTAC	2454
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGluLysLysThr	140
DB	2455	AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGCGCCCGTAAAGCAGAGAAC	2514
QY	141	MetGlyTyrGluLysAsnThrGluArgLeuTyrProArgAspGlyValLeuLysGlu	160
DB	2515	ATGGGCTGGAGGCTTCCACGAGGCGCTTACCCCGGACGGCGGTGTAAGGGCGAG	2574
QY	161	IleHISLysAlaLeuLysLeuLysAspGlyGlyHISLysLeuValGluPheLysSerIle	180
DB	2575	ATCCACAAGGCGCTGTAAGTGAAGGCGGCGGCACTACCTGTGAGTTCAAGTCCATC	2634
QY	181	TyrMetAlaLysLysProValGluLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
DB	2635	TACATGGCCAAAGAGCCGATGACGCTGCCGCTACTACTAGTGAAGTCCCAAGCTGGAC	2694
QY	201	IleThrSerHISAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	2695	ATCACTCTCCCAACGAGCACTACACATCGTGAAGCATGAGGCGCACCGAGGCGCGC	2754
QY	221	HISLysLeuPheLeu	225
DB	2755	CACCACTGTCTCTG	2769

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 12, 2006, 05:49:00 ; Search time 479 Seconds  
(without alignments)  
3130.597 Million cell updates/sec

Title: US-10-006-922A-12

Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFVRMEG.....EDYIVGEYERTGHHFL 225

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODBL=frame+ p2n.model -DEV=xlp  
-O=/cgaz2.1/USPFC.spool.p/US1006922/runat.10012006.162408.18913/app.query.fasta\_1.391  
-DB=N\_Geneseq -QMT=fastcap -SURFIX=p2n.rng -MINMATCH=0.1 -LOOPT=0 -LOOPEXT=0  
-UNIT5=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.coi -LIST=45  
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=PCO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US1006922 @CGEN 1.1 1096 @runat.10012006.162408.18913 -NCPU=6 -ICPU=3  
-NO WMAP -LANG=OTHER -NEG SCORES=0 -WALT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq.21:\*

1:	geneseqn18908:*
2:	geneseqn19908:*
3:	geneseqn20008:*
4:	geneseqn20018:*
5:	geneseqn20028:*
6:	geneseqn20038:*
7:	geneseqn20048:*
8:	geneseqn20058:*
9:	geneseqn20068:*
10:	geneseqn20078:*
11:	geneseqn20088:*
12:	geneseqn20098:*
13:	geneseqn20108:*
14:	geneseqn20118:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	100.0	678	4	AAH47654 Discosoma
2	1214	100.0	678	4	AAD13053 Aadi1053 Discosoma
3	1214	100.0	678	4	AAD11142 Aadi1142 Discosoma
4	1214	100.0	678	6	ABA95905 Yeast opt

5	1214	100.0	678	6	ABA95922	ABA95922 Yeast opt
6	1214	100.0	678	6	ABA95921	ABA95921 Yeast opt
7	1214	100.0	678	6	ABA95920	ABA95920 Yeast opt
8	1214	100.0	678	6	AAD46278	AAD46278 Discosoma
9	1214	100.0	678	6	AAD28207	AAD28207 Discosoma
10	1214	100.0	678	10	ADCC24127	ADCC24127 Discosoma
11	1214	100.0	678	10	ADF70404	ADF70404 Discosoma
12	1214	100.0	678	11	ADL46204	ADL46204 Discosoma
13	1214	100.0	678	11	ADN3978	ADN3978 Wild-type
14	1214	100.0	678	12	ADL36420	ADL36420 Discosoma
15	1214	100.0	678	12	ADM97768	ADM97768 D sp red
16	1214	100.0	695	3	AAA48743	AAA48743 Humanised
17	1214	100.0	859	6	AAL47952	AAL47952 Discosoma
18	1214	100.0	859	7	ADY51715	ADY51715 Discosoma
19	1214	100.0	859	8	AAD53432	AAD53432 Discosoma
20	1214	100.0	859	10	AAD61969	AAD61969 Discosoma
21	1214	100.0	859	14	ADX26533	ADX26533 Discosoma
22	1214	100.0	3111	9	ACA62995	ACA62995 Plasmid D
23	1210	99.7	681	10	ADCC24134	ADCC24134 Discosoma
24	1210	99.7	681	11	ADL46205	ADL46205 Human cod
25	1210	99.7	681	11	ADL46225	ADL46225 Human cod
26	1210	99.7	723	8	ABE22476	ABE22476 Mammalian
27	1210	99.7	1638	10	ADL18131	ADL18131 RFP:PS(NI
28	1210	99.7	1647	10	ADL18155	ADL18155 RFP:PS(HI
29	1210	99.7	2721	13	ADS75466	ADS75466 Fibrohexa
30	1210	99.7	2772	13	ADS75468	ADS75468 Fibrohexa
31	1210	99.7	4692	6	AAL47954	AAL47954 Modified
32	1210	99.7	4692	10	ACC44640	ACC44640 Vector PD
33	1210	99.7	5436	4	AAD10003	AAD10003 Plasmid p
34	1210	99.7	6893	10	ADE24111	ADE24111 Proviral
35	1210	99.7	7910	4	AAD09979	AAD09979 pB1T(dHSP
36	1210	99.7	9320	6	ABS56664	ABS56664 Plasmid p
37	1210	99.7	9658	12	ADL04104	ADL04104 Vector PT
38	1210	99.7	9678	12	ADL04103	ADL04103 Vector PT
39	1210	99.7	10263	12	ADL04102	ADL04102 Vector PT
40	1209	99.6	681	6	ABA95906	ABA95906 Yeast opt
41	1209	99.6	681	6	ABA95907	ABA95907 Yeast opt
42	1207	99.4	678	6	AAD28208	AAD28208 Discosoma
43	1205	99.3	681	4	AAH47656	AAH47656 Anthozoon
44	1203	99.1	7508	13	ADR30814	ADR30814 Zebrafish
45	1197	98.6	6985	10	ACA55359	ACA55359 Transfom

## ALIGNMENTS

RESULT 1  
AAH47654  
ID AAH47654 Bstrand; cDNA; 678 BP.  
XX  
AC AAH47654;  
XX  
DT 30-NOV-2001 (first entry)  
XX  
DE Discosoma sp. red drFP583 protein coding sequence.  
XX  
OS Discosoma sp.  
XX  
PN WO200162919-A1.  
XX  
PD 30-AUG-2001.  
XX  
PF 13-FEB-2001; 2001WO-US004625.  
XX  
PR 23-FEB-2000; 2000US-0184732P.  
XX  
PA (AURO-) AURORA BIOSCIENCES CORP.  
XX  
PI Nelson D, Zamaira E, Tsien R;  
DR WPI; 2001-557704/62.  
XX

Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise functional red fluorescent proteins, and the encoding nucleic acids, with key mutations for improving the proteins function.

Disclosure; Page 83; 90pp; English.

The invention provides a nucleic acid encoding functional red fluorescent CC protein (II) that differs from the sequence of an Anthozoa red fluorescent protein by at least one amino acid substitution, and with different fluorescent properties. The red fluorescent protein of the invention can be expressed by standard recombinant methodology. (II) are used a fluorescent markers and FRET partners. It is used for identifying protein-protein interactions. (II) is also suitable for multiplexed fluorescent analysis and FRET-based applications using existing Aequorea fluorescent proteins. (II) has improved brightness, reduced spectral cross talk, and is rapidly and efficiently expressed in mammalian cells. The key mutations in the encoding nucleic acids provide improved folding, brightness, and create (II) with sharper, more defined excitation and emission peaks when expressed in mammalian cells. The present sequence CC represents the coding sequence of a Discosoma sp. drfp583 protein, an anthozoa fluorescent protein

Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,3e-140 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-006-922a-12 (1-225) x AAHA7654 (1-678)

Qy 1 MetArgSerSerLyAsnValIleLySgluphMetArgPheLyValArgMetGluGly 20  
 Db 1 ATGAGGCTCTCCAGAAATGTTATCAAGAGTTCATGAGGTTCATTAAGCTTCGATGGAAGA 60  
 Qy 21 ThrValaPngLyHleGluPhngLulIeGluGlyGluGlyGluGlyArgProTyrGluGly 40  
 Db 61 ACGGCAATGGGCGACGAGTTTGAAATGAAAGCGGAGGAGGAGGCGCATACGAAAGC 120  
 Qy 41 HisAsnThrValLyLeuLyValThrLySglGlyProLeuProPheAlaTyrAspIle 60  
 Db 121 CACATACCGTAAAGCTTAAGGTACCAAGGGGGAACCTTTGCCATTGCTTGGGATTT 180  
 Qy 61 LeuSerProGlnPhngLntTyrGlySerLyValTyrValLyHisPrcAlaAspIlePro 80  
 Db 181 TTGTCCACCAATTTCAGTATGGAAGCAAGGATATGTCAAGCAACCTGCCGACATACCA 240  
 Qy 81 AspTyrLybLyLeuSerPheProGluGlyPheLySTPrGluArgValMetAsnPhngL 100  
 Db 241 GACTATTAAGCGTCAATTCCTGTAAGGATTTAAATGGGAAAGGTCATGAACTTGA 300  
 Qy 101 AspGlyGlyValIValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 Db 301 GACGGTGGCGTCTGTAAGTCAAGATTCGCAATTTGCGAGATGGCTGTTTCATCTAC 360  
 Qy 121 LySValLybPheIleGlyValaAsnPhngPrcSerAspGlyProValMetGlnLybLyThr 140  
 Db 361 AAGGCAAGTTCATTTGGCGTGAACCTTCCCTCCGATGACCTGTTATCAAAAGAGACA 420  
 Qy 141 MetGlyTTPrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLybGlyGlu 160  
 Db 421 ATGGCGTGGGAAGCGACGATGAGCGTTGTATCTCTGTGATGGGTGTGAAGGAAG 480  
 Qy 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
 Db 481 ATTCAATAGGCTCTGAAGCTGAAGACGCGTGTCTATTACCTAGTTGAATTCAAAAGTATT 540  
 Qy 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValaAspSerLybLeuAsp 200  
 Db 541 TACATGGCAAGAGCGCTGTGCACTACAGGGGTACTACTATGTTGACTCCAAACTGGAT 600

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
 Db 601 ATAAACAGCCACACAGAACTATACATCTTACGACGATTAAGAAACGAGGACGC 660

Qy 221 HisHisLeuPheLeu 225  
 Db 661 CACCATCTGTTCTT 675

RESULT 2

AAD13053  
 ID AAD13053 standard; cDNA; 678 BP.

AC AAD13053;

DT 16-OCT-2001 (first entry)

DE Discosoma sp. "red" anthozoa fluorescent protein, drfp583 cDNA.

KW Protein destabilisation; linker moiety; reporter moiety; disease model;

KW linear multimerised domain; -NH-ubiquitin protein endoprotease;

KW transgenic animal; transgenic plant; disease resistance;

XX anthozoa fluorescent protein; natural fluorescent protein; ss.

OS Discosoma sp.

PN WO200157242-A2.

XX 09-AUG-2001.

PF 02-FEB-2001; 2001WO-US003791.

PR 04-FEB-2000; 2000US-00498098.

XX (AURO-) AURORA BIOSCIENCES CORP.

XX Stack JH, Whitney M, Cubitt AB, Pollok BA;

XX WPI; 2001-488890/53.

XX Destabilizing proteins in living cells, by coupling a target protein to

PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin

PT protein endoproteases, comprising two copies of the domain.

XX Disclosure; Page 110; 171pp; English.

XX The present invention relates to a method for destabilizing a target

CC protein in a cell. The method comprises a linker moiety which operatively

CC couples a target protein (a reporter moiety) to a linear multimerised

CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein

CC endoproteases. The method is useful for detecting an activity such as

CC protease, protein kinase or phosphoprotein phosphatase activity and is

CC also useful for identifying novel modulators of these activities. The method is

CC transnational activators, such as proteolysis, phosphorylation,

CC phosphorylation, glycosylation, methylation, sulfation, prenylation,

CC disulfide bond formation and ADP-ribosylation within cells. The

CC recombinant animal is useful as disease models and transgenic plants with

CC improved disease resistance or other favourable traits. The present

CC sequence is Discosoma sp. "red" anthozoa fluorescent protein, drfp483

CC cDNA which is a natural fluorescent protein used as a reporter moiety in

XX the exemplification of the invention

Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,3e-140 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0



US-10-006-922a-12 (1-225) x AAD13053 (1-678)

QY 1 McrArGSeSerLySaSnValIlleYsgIuphMeArGpHeLySaValArGMeGluGly 20  
 DB 1 ATGAGGCTCTTCCAAAGATGTATCAAGAGGTTCAAGAGTTTAAGGTTTCGATCGAAGGA 60  
 QY 21 ThrValAsnGlyVHlsgIuphEgIuIlegIuGlyGluGlyGluGlyAArgProTyrgIuGly 40  
 DB 61 ACGGTCAATGGGCAAGAGTTGTAATAGAAAGCGAAGAGAGGAGGCGCATACGAAGGC 120  
 QY 41 HlSaSnThrValIlyLeuLySaValThrlYsgIyGlyProLeuProPhelAaTrpAspIle 60  
 DB 121 CACCAATACCGTTAAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTTCGGAGTAAT 180  
 QY 61 LeuSerProGlnPhelGlnTyrgIySerLySaValTyValIlyHlSaProIlaAspIlePro 80  
 DB 181 TTGTCAACCAATTTCAAGTAGAAGCAAGATATGTCAAGCAACCTCCGACATACCA 240  
 QY 81 AspTyLyLyLeuSerPheProGluGlyPheLySTrGluuArgValMetAsnPhelGlu 100  
 DB 241 GACTATTAATAAGCTGTCACTTCCTGAAGGATTTAAATGGGAAAGGCTCATGAACCTTGAA 300  
 QY 101 AspGlyGlyValValThrValThrgIlnAspSerSerLeuGlnAspGlyCySpheIleTy 120  
 DB 301 GACGGTGGCGTCTGTTACTGTAAACCCAGGATTCACAGGATTCAGGATGGCTGTTTCATCTAC 360  
 QY 121 LygValIyPheIleGlyValAsnPhepProSerAspGlyProValMetGlnIySlyThr 140  
 DB 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGAGACCTGTATGCAAAAGAAAGACA 420  
 QY 141 MeIGlyTrpGluIlaSerThrgIlnArgLeuTyProArgAspGlyValIleuYsgIyGlu 160  
 DB 421 ATGGGCTGGGAAGCCAGCACTGAGGCTTTGTATCCCGATGCGGCTGGTGAAGAGAG 480  
 QY 161 IlleHlLyAlaIleuLySaValIyAspGlyVHlYsIyLeuValGluPheLySerIle 180  
 DB 481 ATTCATTAAGGCTCTGAAGCTGAAGAGCGGTGTCACTTACTAGTTAAATTCMAAAGTAT 540  
 QY 181 TyrMetAlaIyLyPheProValGlnLeuProGluTyTrTyTrValAspSerLyLeuAsp 200  
 DB 541 TAAATGGCAAAAGAGCTGTGACGCTACCAAGGTACTACTATGTGATCTCCAAACTGGAT 600  
 QY 201 IlleThSerHlSaSnGluAspTyTrHlIleValIgluGlnTyrgIlnArgThrgIuGlyArg 220  
 DB 601 ATTAACAAGCCACAAAGACATACATGCTTGAGACAGTATGAAGAACCAGGGGAGCG 660  
 QY 221 HlSaHlLeuPheLeu 225  
 DB 661 CACCATCTGTTCCCT 675  
 RESULT 3  
 AAD11142 ID AAD11142 standard; DNA; 678 BP.  
 XX AAD11142;  
 AC XX  
 DT 24-SEP-2001 (first entry)  
 DE Discosoma species Anthozoa fluorescent protein, drfp583 encoding DNA.  
 XX  
 XX Transmembrane potential; biological membrane; fluorescent ion; detection;  
 KM Transmembrane potential; biological membrane; fluorescent ion; detection;  
 KM Transmembrane potential; biological membrane; fluorescent ion; detection;  
 KM Transgenic organism; drfp583 protein; de.  
 XX  
 OS Discosoma sp.  
 XX  
 XX MO200142211-A2.  
 PN  
 PD 14-JUN-2001.  
 PF 12-DEC-2000; 2000MO-US033739.  
 XX

PR 13-DEC-1999; 99US-00459956.  
 XX  
 XX (REGC ) UNIV CALIFORNIA.  
 XX  
 XX Tsien RY, Gonzalez JB;  
 XX  
 DR WPI, 2001-457276/49.  
 XX  
 XX Determining electrical potential across a membrane in biological systems,  
 PT comprising introducing two reagents, exposing the membrane to light and  
 PT measuring the energy transfer.  
 PS  
 PS Disclosure; Page 150; 154pp; English.  
 XX  
 XX The patent discloses optical methods and compositions for determining  
 CC transmembrane potential across biological membranes in living cells. The  
 CC method of determining the electrical potential across a membrane in a  
 CC biological system comprises introduction of two reagents, a first reagent  
 CC comprising a hydrophobic fluorescent ion capable of redistributing from a  
 CC first face of the membrane to a second face of the membrane in response  
 CC to membrane potential change and a second reagent that label the first  
 CC face or the second face of the membrane, which comprises a chromophore  
 CC capable of undergoing energy transfer by either donating or accepting  
 CC excited state energy to the fluorescent ion. The membrane is then exposed  
 CC to excitation light and the energy transfer between the reagents is  
 CC measured and related to the membrane potential. The method is useful for  
 CC detecting changes in membrane potential in subcellular organelle  
 CC membranes in biological systems. The method is used for screening of test  
 CC chemicals for activity to modulate the activity of target ion channel.  
 CC The invention also provides a transgenic organism comprising a first  
 CC reagent that comprises a charged hydrophobic fluorescent molecule and a  
 CC second reagent comprising a bioluminescent or naturally fluorescent  
 CC protein. The present sequence is Discosoma species (red) DNA encoding an  
 CC Anthozoa fluorescent protein (FP), drfp583  
 XX  
 SO Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,3e-140 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-006-922a-12 (1-225) x AAD11142 (1-678)  
 QY 1 McrArGSeSerLySaSnValIlleYsgIuphMeArGpHeLySaValArGMeGluGly 20  
 DB 1 ATGAGGCTCTTCCAAAGATGTATCAAGAGGTTCAAGAGTTTAAGGTTTCGATCGAAGGA 60  
 QY 21 ThrValAsnGlyVHlsgIuphEgIuIlegIuGlyGluGlyGluGlyAArgProTyrgIuGly 40  
 DB 61 ACGGTCAATGGGCAAGAGTTGTAATAGAAAGCGAAGAGAGGAGGCGCATACGAAGGC 120  
 QY 41 HlSaSnThrValIlyLeuLySaValThrlYsgIyGlyProLeuProPhelAaTrpAspIle 60  
 DB 121 CACCAATACCGTTAAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTTCGGAGTAAT 180  
 QY 61 LeuSerProGlnPhelGlnTyrgIySerLySaValTyValIlyHlSaProIlaAspIlePro 80  
 DB 181 TTGTCAACCAATTTCAAGTAGAAGCAAGATATGTCAAGCAACCTCCGACATACCA 240  
 QY 81 AspTyLyLyLeuSerPheProGluGlyPheLySTrGluuArgValMetAsnPhelGlu 100  
 DB 241 GACTATTAATAAGCTGTCACTTCCTGAAGGATTTAAATGGGAAAGGCTCATGAACCTTGAA 300  
 QY 101 AspGlyGlyValValThrValThrgIlnAspSerSerLeuGlnAspGlyCySpheIleTy 120  
 DB 301 GACGGTGGCGTCTGTTACTGTAAACCCAGGATTCACAGGATTCAGGATGGCTGTTTCATCTAC 360  
 QY 121 LygValIyPheIleGlyValAsnPhepProSerAspGlyProValMetGlnIySlyThr 140

```

Db      361 AAGTCAGATTCATTGGCGTGAACCTTCCGATGACCTGTATGCAAAAGAGACA 420
Qy      141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProAlaGlyValLeuLeuGlyGlu 160
Db      421 ATGGCTGGGAAGCCAGACGTGAGGTTGTATCCTCGTAGAGGGGTGTGAAGAGAG 480
Qy      161 IleHisValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
Db      481 ATTCAATAGGCTCTAGAGCTGAAAGACGGTGTCTATTACCTAGTTGAATTCAAAGTATT 540
Qy      181 TyrMetAlaValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 200
Db      541 TACATGGCAAGAAAGACCTGTGACCTACAGGGGTACTACTATGACTCCAAATCGGAT 600
Qy      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      601 ATTAAGAACCCACAAAGAAAGACTATCAATCGTTGAGAGTGAAGAAAGAGGAGAGC 660
Qy      221 HisHisLeuPheLeu 225
Db      661 CACCATCTGTTCTT 675

RESULT 4
ABA95905
ID      ABA95905 standard; cDNA; 678 BP.
AC      ABA95905;
DT      29-MAY-2002 (first entry)
DE      Yeast optimised RFP encoding cDNA SEQ ID NO 1.
XX      Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;
KM      Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;
KW      Escherichia coli; green fluorescent protein; biotechnology; gene; ss.
XX      Anthozoa.
FH      Key      Location/Qualifiers
FT      CDS      1..678
FT      /tag=      a
FT      /product= "Yeast optimised Red Fluorescent Protein"
XX      DE20001395-U1.
XX      PD      15-MAR-2001.
XX      PF      27-JAN-2000; 2000DE-02001395.
XX      PR      27-JAN-2000; 2000DE-02001395.
XX      PA      (GPCB-) GPC BIOTECH AG.
XX      WP1; 2002-228394/29.
XX      DR      P-PSDB; ABB08834.
XX      PT      New DNA encoding red fluorescent protein, useful as marker in
XX      biotechnology, has sequence optimized for expression in eukaryotes,
XX      especially yeast or plants.
XX      PS      Claim 1; Fig 1; 19pp; German.
XX      CC      The invention relates to DNA (1) containing either sequence ABA95905 or
XX      sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein
XX      (yRFP). (1) are used to express red fluorescent protein (RFP) in
XX      eukaryotes, especially yeast, especially Saccharomyces cerevisiae and
XX      plants, especially dicotyledonous plants including Nicotiana tabacum or
XX      Arabidopsis thaliana and also in prokaryotes, especially bacteria,
XX      especially Escherichia coli. RFP is useful in the same way as green
XX      fluorescent protein but is more generally applicable in modern
XX      biotechnology. (1) are optimised for expression in yeast and so generate
XX      RFP at higher levels with stronger fluorescence and thus lowers the
XX      detection limit and gives a better signal-to-noise ratio
CG

```

```

XX      SQ      Sequence 678 BP; 198 A; 147 C; 159 G; 174 T; 0 U; 0 Other;
XX      Alignment Scores:
XX      Pred. No.:      2.3e-140
XX      Score:      1214.00
XX      Percent Similarity:      100.00%
XX      Best Local Similarity:      100.00%
XX      Query Match:      100.00%
XX      DB:      6
XX      Gaps:      0
XX      US-10-006-922a-12 (1-225) x ABA95905 (1-678)
Qy      1 MetArgSerSerLeuAsnValIleLeuGluPheMetArgPheLeuValArgMetGluGly 20
Db      1 ATGAGATCTTCTTAAGAACGTCATCAAGGAATTCATGAGATTCAGATTCAAGTTAGAGAGT 60
Qy      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
Db      61 ACTGTTAACGGTCAAGATTGGAATCGAAGTGAAGGTGAAGGTAGACCATACGAAGGT 120
Qy      41 HisAsnThrValLeuLeuLeuLeuValThrLeuGlyGlyProLeuProPheLeuAspIle 60
Db      121 CACAACACTGTCAAGTTGAAGGTTACTAAGGTGTCCATTGCCATTGCGCTTGGAGACATC 180
Qy      61 LeuSerProGluPheGlnTyrGlySerLeuValTyrValLeuHisProAlaAspIlePro 80
Db      181 TTGTCTCAACATTCCAATACGATTCCTTAAGGTTCAAGTCAAGCAACCACTGACATTC 240
Qy      81 AspTyrLeuLeuLeuSerPheProGluGlyPheLeuTyrProGluArgValMetAsnPheGlu 100
Db      241 GACTACAAAGAAAGTGTCTCTCCAGAAAGTTTCAAGTGGGAAAGATCATGAACTTCGAA 300
Qy      101 AspGlyGlyValValThrValThrGluAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      301 GACCGTGTGTGTACTGTTACTCAAGACTCCCTCTTGCAAGCGGTGTGTTCATCTAC 360
Qy      121 LeuValLeuPheIleGlyValAsnPheProSerAspGlyProValMetGlnLeuLeuThr 140
Db      361 AAGGTCAAGTTCACTGGTGTCACTTCCATCTGACGTCACAGTCATGCAAAAGAGACT 420
Qy      141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProAlaGlyValLeuLeuGlyGlu 160
Db      421 ATGGCTGGGAAGCTTACCGAAGCTTTGTATCCCAAGAGACGGTGTGAAGGAGAA 480
Qy      161 IleHisValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
Db      481 ATCCCAAGGCTTGAAGTTGAAGACGGTGTCTACTTGTGCAATTCAAGTCTATC 540
Qy      181 TyrMetAlaValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 200
Db      541 TACATGGCTAAGAACCAAGTCCAAATTCAGGTTACTACTACGTTGACTCTAAGTTGAC 600
Qy      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      601 ATCACTCTTCAAGCAAGAAAGACTACATATCGAACAAATCGAACTGAAAGGTAGA 660
Qy      221 HisHisLeuPheLeu 225
Db      661 CACCATCTGTTCTT 675

RESULT 5
ABA95922
ID      ABA95922 standard; DNA; 678 BP.
AC      ABA95922;
DT      29-MAY-2002 (first entry)
DE      Yeast optimised RFP related DNA SEQ ID NO 19.
XX      Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;
KM      Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;
KW

```

KW Escherichia coli; green fluorescent protein; biotechnology; gene; ss.  
 XX Anthozoa.  
 XX Key Location/Qualifiers  
 FT CDS 1..678  
 PT /tag= a  
 /product= "yeast optimised Red Fluorescent Protein"  
 XX DE20001395-U1.  
 XX PD 15-MAR-2001.  
 XX PF 27-JAN-2000; 2000DE-02001395.  
 XX PR 27-JAN-2000; 2000DE-02001395.  
 XX (GPCB-) GPC BIOTECH AG.  
 XX WPI; 2002-228394/29.  
 XX DR New DNA encoding red fluorescent protein, useful as marker in  
 XX PT biotechnology, has sequence optimized for expression in eukaryotes,  
 XX PT especially yeast or plants.  
 XX PS Disclosure; Page 14-15; 19pp; German.  
 XX CC The invention relates to DNA (I) containing either sequence ABA95905 or  
 CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
 CC (YRFP). (I) are used to express red fluorescent protein (RFP) in  
 CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
 CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
 CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
 CC especially Escherichia coli. RFP is useful in the same way as green  
 CC fluorescent protein but is more generally applicable in modern  
 CC biotechnology. (I) are optimised for expression in yeast and so generate  
 CC RFP at higher levels with stronger fluorescence and thus lowers the  
 CC detection limit and gives a better signal-to-noise ratio. The present  
 CC sequence is that of a polynucleotide encoding the yeast optimised RFP,  
 CC useful to the invention  
 XX SQ Sequence 678 BP; 202 A; 118 C; 159 G; 199 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,3e-140 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0  
 US-10-006-922a-12 (1-225) x ABA95922 (1-678)  
 QY 1 MetArgSerSerIysAsnValIleIyGluPheMetArgPheIysValArgMetGluGly 20  
 DB 1 ATGAGATCTTCTAAGACGTTATTAAGCAATTCATGAGATTCAAGCTTAGAATGGAAGGT 60  
 QY 21 ThrValAsnGluHisGluPheGluIleGluGlyGluGlyArgProIyTrpGluGly 40  
 DB 61 ACTGTTAACGGTCCGATTCGAAATTCGAAAGTGAAGGTAGACCAATTCGAAAGGT 120  
 QY 41 HisAsnThrValIyLeuIyValThrIyGlyGlyProIeuProPheAlaTrpAspIle 60  
 DB 121 CACCAACCTGTTAAGTTCAGAGTTACTTAAGGTCGTCATTCGCAATTCGCTGGGACATT 180  
 QY 61 LeuSerProGluInpGluIyTrpIySerIyValIyValIyHisProAlaAspIlePro 80  
 DB 181 TTGCTCCACCAATTCACATACGGTTCTTAAGCTTACGTTAAGCAACCAAGCTGACATTC 240  
 QY 81 AspTrpIyIyLeuSerPheProGluGlyPheIyTrpGluIyValMetAsnPheGlu 100  
 DB 241 GACTACAAAGAAAGTTCTTTCCTCCAGAAAGTTCAAGTGGGAAAGATTATGAACTTCGAA 300

QY 101 AspGluValValThrValThrGluAspSerLeuGluAspGlyCysPheIleTyr 120  
 DB 301 GACGGTGGTGTGTACTGTTACTGACCTGTTCTTTCGAAAGCGTTGTTACTTTAC 360  
 QY 121 LysValIyPheIleGlyValIAsnPheProSerAspGlyProValMetGluIyIyThr 140  
 DB 361 AAGGTTAAGTTCATTCGTTTACTTCCCATCTGACGCTCCAGTTACCAAGAAAGAACT 420  
 QY 141 MetGlyTrpGluIyAsnThrGluArgLeuTyrProArgAspGlyValIleIyGlyGlu 160  
 DB 421 ATGGTTGGAAAGCTTCTACTGAAAGATTGTACCAAGACGCGTGTGTTGAAAGGTGAA 480  
 QY 161 IleHisIyValIleIyLeuIyLeuIyAspGlyIyHisTyrIleValIyGluPheIySerIle 180  
 DB 481 ATTCACAAAGCTTTAAGTTGAAGACGCGTGTCTACTCTGTTGTAATTCAAGCTTATT 540  
 QY 181 TyrMetAlaIyIyProValIyGluIyProGlyIyTrpIyTrpValIyAspSerIyLeuAsp 200  
 DB 541 TACATGCTTAGAACACCACTGCAATTCGCAAGTTACTACTGACTGACTTAAGTTGAC 600  
 QY 201 IleThrSerHisAsnGluAspTyrThrIleValIyGluIyTrpGluArgThrGluIyArg 220  
 DB 601 ATTACTTCTTCACAAAGAACTACACTATTGTTGAACATACGAAAGAACTGAAGTAGA 660  
 QY 221 HisHisIeuPheIeu 225  
 DB 661 CACCACTTGTTCTTG 675  
 DB RESULT 6  
 ABA95921  
 ID ABA95921 standard; DNA; 678 BP.  
 XX ABA95921;  
 XX 29-MAY-2002 (first entry)  
 XX  
 XX Yeast optimised RFP related DNA SEQ ID NO 18.  
 XX  
 XX Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
 XX Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
 XX Escherichia coli; green fluorescent protein; biotechnology; gene; ss.  
 XX Anthozoa.  
 XX Key Location/Qualifiers  
 FT CDS 1..678  
 FT /tag= a  
 FT /product= "yeast optimised Red Fluorescent Protein"  
 XX DE20001395-U1.  
 XX PD 15-MAR-2001.  
 XX PF 27-JAN-2000; 2000DE-02001395.  
 XX PR 27-JAN-2000; 2000DE-02001395.  
 XX (GPCB-) GPC BIOTECH AG.  
 XX WPI; 2002-228394/29.  
 XX P-PSDB; ABB0834.  
 XX New DNA encoding red fluorescent protein, useful as marker in  
 XX PT biotechnology, has sequence optimized for expression in eukaryotes,  
 XX PT especially yeast or plants.  
 XX PS Disclosure; Page 14; 19pp; German.  
 XX CC The invention relates to DNA (I) containing either sequence ABA95905 or  
 CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
 CC (YRFP). (I) are used to express red fluorescent protein (RFP) in  
 CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
 CC plants, especially dicotyledonous plants including Nicotiana tabacum or

CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of a polynucleotide encoding the yeast optimised RFP,  
CC useful to the invention

XX Sequence 678 BP, 205 A, 129 C, 178 G, 166 T, 0 U, 0 Other;

#### Alignment Scores:

Pred. No.:	2.3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x ABA95921 (1-678)

```
QY      1  MchArGSeSerLyAsnValIleLyGluPheMetArGPhelyValArgMetGluGly  20
        |||
DB      1  ATGAGGCTTCAGAGAAATGTTATCAAGAGATTCAAGAGTTTAAGGTCGCAATGAGAAAGA  60
QY      21  ThrValaAsnGlyHisGluPheGluIleGluGlyGluGlyArgProGlyGluGly  40
        |||
DB      61  ACGGCAATGGGACGAGTTGAAATAGAAAGCGAAGAGAGGGAGGCAATCGAAGGC  120
QY      41  HisAsnThrValLyLeuLyAsnValThrLyGlyGlyProLeuProPheAlaTrpAspIle  60
        |||
DB     121  CACATATACGTTAAAGCTTAAAGGTAAACAAGGGGAGCCTTTGCCATTTGCTTGGGATATT  180
QY      61  LeuSerProGlnPheGlnIleGlySerLyValTyValLyHisProAlaAspIlePro  80
        |||
DB     181  TTGTCACCAACAATTCAGTATGGAAGCAAGGTATGTCAAGCAACCCGCGACATACCA  240
QY      81  AspTyLyLyLeuSerPheProGluGlyPheLySTPGLuATGValMetAsnPheGlu  100
        |||
DB     241  GACTATTAAGAGCTCTCATTTCTCTTAAGGATTTAAATGGAAAGGCTCATGAACCTTTGAA  300
QY     101  AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTy  120
        |||
DB     301  GACGGTGGCGTGTACTGTAACCAAGATTCCAGTTTGCAGGATGGCTGTTTCACTTAC  360
QY     121  LySTValLyPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySerThr  140
        |||
DB     361  AAGTCAAAGTTCATTGGCGGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAGACA  420
QY     141  MetGlyTPGluAlaSerThrGluArgLeuTyProArGAspGlyValLeuLyGlyGlu  160
        |||
DB     421  ATGGGCTGGAAAGCCAGCACTGAAGCGTTTGTATCTCTGTGAATGGGCTTTGAAGGAGAG  480
QY     161  IleHisLyAlaLeuLyLeuLyAsnAspGlyGlyHisTyLeuValGluPheLySerIle  180
        |||
DB     481  ATTCAATTAAGGCTCTGAAGCTGAAGAAGCGTGTCTATTCTTAAGTTGAATTAAGATATT  540
QY     181  TyrMetAlaLyLyPProValGlnLeuProGlyTyTyTyTyValAspSerLyLeuAsp  200
        |||
DB     541  TACATGGCAAGAAAGCCCTGTGACGCTACAGGGTCTACTATGTTGACTCCAAACTGAT  600
QY     201  IleThrSerHisAsnGluAspTyThrIleValGlnGlnTyArgLyArgThrGluGlyArg  220
        |||
DB     601  ATTAACAAGCCACAAGAAAGACTATACATCGTTGAGCATATGAAGAAGACGAGGAGACGC  660
QY     221  HisHisLeuPheLeu  225
        |||
DB     661  CACCATCTGTTCTT  675
QY     661  CACCATCTGTTCTT  675
```

RESULT 7  
ID ABA95920 standard; RNA; 678 BP.  
XX

```
AC      ABA95920;
XX      29-MAY-2002 (first entry)
DT
XX      Yeast optimised RFP encoding RNA SEQ ID NO 16.
DE
XX      Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;
XX      Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;
XX      Escherichia coli; green fluorescent protein; biotechnology; ss.
XX      Anthozoa.
OS
XX      DE20001395-UI.
XX      15-MAR-2001.
XX      27-JAN-2000; 2000DE-02001395.
XX      27-JAN-2000; 2000DE-02001395.
XX      27-JAN-2000; 2000DE-02001395.
XX      (GPCB-) GPC BIOTECH AG.
XX      WPI; 2002-228394/29.
XX      New DNA encoding red fluorescent protein, useful as marker in
XX      biotechnology, has sequence optimized for expression in eukaryotes,
XX      especially yeast or plants.
XX      Disclosure; Page 13; 19pp; German.
XX
XX      The invention relates to DNA (I) containing either sequence ABA95905 or
XX      sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein
XX      (yRFP). (I) are used to express red fluorescent protein (RFP) in
XX      eukaryotes, especially yeast, especially Saccharomyces cerevisiae and
XX      plants, especially dicotyledonous plants including Nicotiana tabacum or
XX      Arabidopsis thaliana and also in prokaryotes, especially bacteria,
XX      especially Escherichia coli. RFP is useful in the same way as green
XX      fluorescent protein but is more generally applicable in modern
XX      biotechnology. (I) are optimised for expression in yeast and so generate
XX      RFP at higher levels with stronger fluorescence and thus lowers the
XX      detection limit and gives a better signal-to-noise ratio. The present
XX      sequence is that of an RNA sequence corresponding to the yeast optimised
XX      RFP encoding cDNA given as SEQ ID NO 1 (ABA95905)
XX
SQ      Sequence 678 BP, 198 A, 147 C, 159 G, 0 T, 174 U, 0 Other;

Alignment Scores:
Pred. No.:      2.3e-140      Length:      678
Score:          1214.00      Matches:      225
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6           Gaps:       0

US-10-006-922A-12 (1-225) x ABA95920 (1-678)
QY      1  MchArGSeSerLyAsnValIleLyGluPheMetArGPhelyValArgMetGluGly  20
        |||
DB      1  AUGAGACUUCUAAAGAGCAUCACAAAGAAUUCUGAGAUCCAAAGGUAAGAAUUGGAGAGU  60
QY      21  ThrValaAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyGluGly  40
        |||
DB      61  ACUGUUAACGGUCACGAUUCGAAUUCGAAGUGAAGUUAAGUUAACCAUACGAAGU  120
QY      41  HisAsnThrValLyLeuLyAsnValThrLyGlyGlyProLeuProPheAlaTrpAspIle  60
        |||
DB     121  CACAAACACUGUCAAAGUUGAAGUUAACUAAAGGUGUCCAUUGCCAUUGCGUUGGACAUC  180
QY      61  LeuSerProGlnPheGlnIleGlySerLyValTyTyTyValLyHisProAlaAspIlePro  80
        |||
DB     181  UUGUCUCCACAUAUCCAAUACGGUUCUAAAGUUCUACGUAAGCAACCCAGCUCAUAUCCA  240
QY      81  AspTyLyLyLeuSerPheProGluGlyPheLySTPGLuArgValMetAsnPheGlu  100
```

DB 241 GACUACAAAGAGUUCUCCAGAGGUGUUCAGUGGAGAAAGACUAGACUCCGAA 300  
 QY 101 AARPGLVGLVValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 DB 301 GACGGGUGUGUGUUCUUCUUAACUACAGACUCCUUGGCAAGACGGUGUGUUCUUC 360  
 QY 121 LysValIlePheIleGlyValIlePhePheProSerAspGlyProValMetGlnIleValThr 140  
 DB 361 MAGGUCUAUUCUUGGUGUUCACUCCACUUGACGGUCACGUCUAGCAAAAGAGACU 420  
 QY 141 MetGlyTyrGlnIleAspSerThrGluArgLeuTyrProArgAspGlyValIleIleValGlu 160  
 DB 421 AUGGUGGAGAGUCUUCACCGAACGUGUUAACCAAGAGACGGUGUCUAGGGUGAA 480  
 QY 161 ILeHleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 180  
 DB 481 AUCGACAAAGCCUUGAGUUGAGAGACGGUGUCACUACUUGGUCAGAUUCAGUCUAC 540  
 QY 181 TyrMetAlaIleValIleProValGlnLeuProGlyTyrTyrTyrValIleAspSerIleAsp 200  
 DB 541 UACAGGCUAAGAGACGACGACUUGCCAGUACGUCUACUACUUGACUUAAGUGAGAC 600  
 QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnIleGlnIleValArgThrGlnIle 220  
 DB 601 AUCACCUUCACAAAGAGACUACUACUUGGAGAACAAUACGAAAGUACUAGAGUAG 660  
 QY 221 HisHleIleuPheIleu 225  
 DB 661 CACCAUCUUGUCUG 675  
 RESULT 8  
 AAD46278  
 ID AAD46278 standard; DNA; 678 BP.  
 AC AAD46278;  
 DT 27-DEC-2002 (first entry)  
 XX  
 DE Discosoma sp. drfp583 (NFP-6) wild-type protein encoding DNA.  
 XX  
 KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
 KW fluorescence activated cell sorting application; fluorescent timer;  
 KW biosensor; fluorescence resonance energy transfer application; FRET;  
 KW colouring agent; recombinant DNA application; analyte detection assay;  
 KW biosensor; second messenger detector; drfp583 protein; NFP-6; gene; ds.  
 XX  
 OS Discosoma sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..678  
 FT /tag= a  
 FT /product= "drfp583 wild-type protein"  
 XX  
 PN MO200268459-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 20-FEB-2002; 2002MO-US005749.  
 XX  
 PR 21-FEB-2001; 2001US-0270983P.  
 PR 04-DEC-2001; 2001US-00006922.  
 XX  
 PA (CLON-) CLONTECH LAB INC.  
 XX  
 PI Lukyanov S, Lukyanov K, Yanushkevich Y, Savitsky A, Fradkov A;  
 XX  
 DR WPI; 2002-691654/74.  
 DR P-PSDB; AAE28833.  
 XX  
 PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
 PT of an aggregating Chladerian chromo- or fluorescent protein or mutant for  
 PT analyte detection assays or fluorescence activated cell sorting

PT applications.  
 XX  
 XX Disclosure; Page 70; 80pp; English.  
 PS  
 CC The invention relates to nucleic acid molecules encoding non-aggregating  
 CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
 CC useful in analyte detection assays, as colouring agents, as markers in  
 CC recombinant DNA applications, as biosensors or filters, in fluorescence  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, in screening assays, as second  
 CC messenger detectors, in fluorescence activated cell sorting applications,  
 CC in protease cleavage assays or as fluorescent timers. The present  
 CC sequence is a DNA encoding Discosoma sp. drfp583 (NFP-6) wild-type  
 CC protein of the invention  
 XX  
 SQ Sequence 678 BP; 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,3e-140 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-006-922A-12 (1-225) x AAD46278 (1-678)  
 QY 1 MetArgSerSerIleAsnValIleIleValGluPheMetArgPheIleValArgMetGlnIle 20  
 DB 1 ATGCGCTCTCCAAAGACGTCATCAAGAGGTTCAATCGCTTCAAGGTGCGCATGAGGGC 60  
 QY 21 ThrValAsnGlyHisGluPheGluIleGluIleGluIleGluIleValArgProTyrGlnIle 40  
 DB 61 ACCGTGAACGGCCACAGATTCCAGATCGAGGGCGAGGGGCGAGGGCCGCTACAGAGGGC 120  
 QY 41 HisAsnThrValIleIleIleValIleThrIleValIleValIleProIlePheIleValIle 60  
 DB 121 CACAAACACCGTAGAGGTGAAGGACCAAGGCGGCGCCCTCGCTCCCTGGGACATC 180  
 QY 61 LeuSerProGlnIlePheGlnIleTyrGlySerIleValIleValIleAsnIlePro 80  
 DB 181 CTCTCCCCCAAGTTCAGTAGACGCTCCAGAGGTGTAGTAGAGACCCCGCGACATCCCC 240  
 QY 81 AspTyrIleValIleIleSerPheProGluIlePheIleTyrGlnIleValMetAsnPheGln 100  
 DB 241 GACTACAAAGAGCGTGTCTTCCCGAGGGCTTCAGATGGAGCGCGTAGTAACCTTCGAG 300  
 QY 101 AARPGLVGLVValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 DB 301 GACGGGUGUGUGUUCUUCUUAACUACAGACUCCUUGGCAAGACGGUGUGUUCUUC 360  
 QY 121 LysValIlePheIleGlyValIlePhePheProSerAspGlyProValMetGlnIleValThr 140  
 DB 361 MAGGUCUAUUCUUGGUGUUCACUCCACUUGACGGUCACGUCUAGCAAAAGAGACU 420  
 QY 141 MetGlyTyrGlnIleAspSerThrGluArgLeuTyrProArgAspGlyValIleIleValGlu 160  
 DB 421 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGACGCGTGTGAAGGGCGAG 480  
 QY 161 ILeHleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 180  
 DB 481 ATCCACAAAGCCCTGAAGCTGAAGAGACGGGCGGCACTCTGTGTGAGTTCAAGTTCATC 540  
 QY 181 TyrMetAlaIleValIleProValGlnLeuProGlyTyrTyrTyrValIleAspSerIleAsp 200  
 DB 541 TACATGGCCAAAGAGCCCGTGACGCTGCCGGGTACTAGTAGTGACTCAAGCTGAC 600  
 QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnIleGlnIleValArgThrGlnIle 220  
 DB 601 ATCACTCCCAAGAGAGAGACTACACCATGTGAGAGAGACGAGGCGAGGCGCGC 660  
 QY 221 HisHleIleuPheIleu 225

Db 661 CACCACCTGTTCCRG 675

RESULT 9

AD28207

AD28207 standard; DNA; 678 BP.

XX

AC AD28207;

XX

22-APR-2002 (first entry)

XX

De Diacosoma sp. humanised wild-type Anthozoa protein drFP583 DNA.

XX

KW Fluorescent timer protein; protein movement; translocation; trafficking;

KM promoter activity; gene expression; transgenic plant; gene modification;

KM protein age; anthozoa protein; drFP583; ds.

XX

OS Diacosoma sp.

XX

Key Location/Qualifiers

FT CDS 1..678

FT /tag= a

FT /product= "Humanised wild-type Anthozoa protein drFP583"

XX

MO200196373-A2.

XX

20-DEC-2001.

XX

13-JUN-2001; 2001WO-US019097.

XX

14-JUN-2000; 2000US-0211607P.

XX

(CLON-) CLONTECH LAB INC.

XX

Pradkov AF, Terakikh A;

XX

WPI; 2002-154595/20.

XX

P-PSDB; AAE17540.

DR

XX

New fluorescent timer proteins comprising an emission spectrum that

FT changes over time from a first wavelength to a second wavelength, useful

FT for monitoring intracellular protein movement, translocation, trafficking

FT or stability.

XX

Example 1; Fig 1; 89pp; English.

XX

The invention relates to a fluorescent timer protein having an emission

CC spectrum that changes over time after synthesis from a first wavelength

CC to a second wavelength. The fluorescent timer proteins are useful in

CC monitoring the activity of a promoter, determining the age of a protein,

CC identifying an agent that modulates the activity of a promoter and in

CC enriching a population of cells comprising a fluorescent timer protein.

CC The fluorescent timer proteins are also useful for assessing gene

CC expression during development of a multicellular organism or during

CC cellular differentiation, in response to a drug or other inducer of

CC promoter activity, as a reporter to serve as a read-out of promoter

CC activity, monitoring intracellular protein movement or translocation,

CC protein trafficking, or protein stability, to investigate temporal

CC aspects of the activity of a regulatory element, for determining cell

CC fate during development and organ remodelling, in spatial and temporal

CC visualisation of newly synthesised proteins and accumulated proteins, e.g.

CC in distinguishing between newly formed and pre-existing structures, e.g.

CC membrane junctions and extracellular matrix components. The fluorescent

CC timer proteins may further be used to investigate mechanisms where photobleaching

CC techniques are employed, as detectable labels, as selectable markers, as

CC biosensors in prokaryotic and eukaryotic cells, in protease cleavage

CC assays, and as second messenger detectors. The nucleic acids can be used

CC to generate transgenic, non-human plants or animals or site-specific gene

CC modifications in cell lines. The present sequence is a DNA encoding

CC Diacosoma sp. humanised wild-type Anthozoa protein drFP583 used for

CC generating fluorescent proteins

XX

Sequence 678 BP; 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;

XX

5Q

## Alignment Scores:

Pred. No.:	2,3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922a-12 (1-225) x AD28207 (1-678)

QY	1	MetArgSerSerIysAenValIleIleGluPheMetArgPheIysValIArgMetGluGly	20
Db	1	ATGGCGTCTCTCCAAAGACGTGATCAAGAGTTCATGCGCTTCAAGGTGGCATGAGGGCC	60
QY	21	ThrValMetGlyHisGluPheGluIleGluGluGlyValGluGlyIleArgProTyrGluGly	40
Db	61	ACCGTGAACGGCCACGAGTTCAGATGAGGCGCAGGAGCGCCGCCCTACGAGGAC	120
QY	41	HisAsnThrValIysLeuIysValThrIleGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAACACCGTGAAGCTGAAGTGAACCAAGGCGGCCCTGCGCTTGCGGAGATC	180
QY	61	LeuSerProGluPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro	80
Db	181	CTGTCCCCCGAGTTCAGTACGCGCTCCAGGTGATGAGGACCCCGCGACATCCCC	240
QY	81	AspTyrIleIysLeuSerPheProGluGlyPheIleTrpGluArgValMetAsnPheGlu	100
Db	241	GACTACAAAGAGCGTCTCTCCCGAGGCGCTTCAAGTGGAGCGCGTATGATCACTTCAG	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGCGCGCGTGGAGCGGTGACCCAGACCTCCCTCGACAGACGGGTGCTTCAATCAC	360
QY	121	LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIleLysThr	140
Db	361	AAGGTGAAGTTCATCGCGCTGGAATCTCCCTCCACGCGCCCGGTGATGCAGAAAGACC	420
QY	141	MetGlyTrpGluIleAsnThrGluArgLeuTyrProArgAspGlyValIleLysGlyIle	160
Db	421	ATGGGCTGGAGGCTTCACCGAGCGCTGTACCCCGCAGCGCGTGTGAAGGGCCAG	480
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACATCGGTGGAGTTCAATCCATC	540
QY	181	TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCCAAAGACCCGTCGAGCTCCGCGCTACTACGTGACCTCCAGCTGGAC	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATCACCTCCCAACGAGACTACACATGTGGAGCATGACGAGCCGACCGGCGC	660
QY	221	HisHisLeuPheLeu 225	
Db	661	CACCACCTGTTCCRG 675	

RESULT 10

ID AD24127

AD24127 standard; DNA; 678 BP.

XX

AC AD24127;

XX

18-DEC-2003 (first entry)

XX

Diacosoma wild-type red fluorescent protein DNA #1.

XX

KW Diacosoma red fluorescent protein; DsRed; AB interface; AC interface;

KM fluorescence energy resonance transfer; FRRT; protein kinase;

KM protein phosphatase; ion indicator; ds.

XX



DR P-PSDB; ADF70403.  
 XX Transformation of cells with a fusion protein of an orphan receptor  
 PT protein with a fluorescent protein useful for identification of ligands  
 FT to the orphan receptor.  
 XX  
 XX  
 PS Disclosure; SEQ ID NO 27; 594bp; Japanese.  
 CC This invention relates to a novel method of identifying ligands to an  
 CC orphan receptor protein which comprises transforming cells with DNA  
 CC encoding a fusion protein of the orphan receptor with a fluorescent  
 CC protein, so that the fusion protein is expressed in the cells (or cell  
 CC membranes isolated from them) and contacting the cells with the potential  
 CC ligand to be tested. A suitable fluorescent protein for incorporation in  
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
 CC identification of ligands binding to an orphan receptor protein.  
 XX  
 SQ Sequence 678 BP; 204 A; 129 C; 179 G; 166 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.3e-140 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-006-922a-12 (1-225) x ADF70404 (1-678)  
 QY 1 MetATGSSerSeryAsnValIlelysgluPhemeCArgphelyValArgMetGluGly 20  
 DB 1 ATGAGGTCTTCCAGAAATCTTATCAAGAGATTCATGAGGTTTAAGGTTCCGATGSAAGA 60  
 QY 21 ThrValaanglyHiegluPhegluilegluglygluglyYargProtyrGluGly 40  
 DB 61 ACGGCAATGGGCGACGAGTTTGAATAGAGCGAAGGAGGAGGCGCATACGAAAGGC 120  
 QY 41 HisaenThrValyleuLeuValThrltysglyglYProleuProPhaAlaTrrAspIle 60  
 DB 121 CACATATACCGTAAAGCTTAAAGGTAAACAAAGGGGAGACCTTGGCCATTTGGGATATT 180  
 QY 61 LeuSerProGluPhenGluNtyrGlySeryValTyVallyshAsProAlaAspIlePro 80  
 DB 181 TTGTACCAACAATTTCAAGTATGSAAGCAAGATATGTCAAGCAACCTGCCGACATACCA 240  
 QY 81 AspTyrlYsLysLeuSerPheProGluGlyPheLystrGluArgValMetAsnPhenGlu 100  
 DB 241 GACATTAATAAAAGCTGTCAATTCCTGAAGGATTTAATGGAAAGGCGATGAACTTTGAA 300  
 QY 101 ArgGlyglYValValThrlValThrlnaPserSerleuGlnaAspGlyCyPheIleTyr 120  
 DB 301 GACGGGTGGCGTCTGTAACGTAAACCAAGATTCGATTTGCGAGATGGCTGTTCACTTAC 360  
 QY 121 LysVallyshPheIleGlyValaAsnPherProSerAspGlyProValMetGlnLysIleThr 140  
 DB 361 AAGGCAAGTTCATTTGGCGGTAAACTTCTCCGATGACCTGTTATGCAAAAGAAAGCA 420  
 QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrrProArgAspGlyValleuLygGlu 160  
 DB 421 ATGGGCTGGGAAGCAGACATGAGCGTTGTATCTCTCGATGGCGGTGGAAGAGAGAG 480  
 QY 161 IleHslslyAlaLeuLysLeuLysAspGlyglYHslsTyrlleuValGluPhelysSerIle 180  
 DB 481 ATTCAATAAGGCTCTGAAAGCTGAAAGACGATGCTATTCCTAAGTTGAATTCAAAAGTATT 540  
 QY 181 TyrMetAlaLysLysProValGluLeuPProGlyTyrrTyrrTyrrValaAspSeryLeuAsp 200  
 DB 541 TACATGGCAAGAAAGCCTGTGCAAGTACCAAGGATCTATGTTGATCTCCAAATCGAT 600  
 QY 201 IleThrserrHlsAsnGluAspTyrrThrlleValGluGlnTyrgluArgThrlGluGlyArg 220  
 DB 601 ATTAACAAGCCACAAGAAAGATATACATCGTTGAGCGATATGAAGAAGACCGAGGAGCGC 660

QY 221 HisHslsleuPhelyeu 225  
 DB 661 CACCATCTGTTCTT 675  
 RESULT 12  
 ID ADL46204  
 ADL46204 standard; DNA; 678 BP.  
 XX  
 AC ADL46204;  
 XX  
 DT 20-MAY-2004 (first entry)  
 DE Discosoma red fluorescent protein (DRed) coding sequence.  
 XX  
 KW ds; gene; red fluorescent protein; DRed; fluorescence; red wavelength;  
 KM oligomerization; tetramerization; immunoassay; hybridization assay.  
 XX  
 OS Discosoma sp.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 1..678  
 FT /\*cag= a  
 FT /product= "DRed protein"  
 XX  
 XX MO2003086446-A1.  
 XX  
 PD 23-OCT-2003.  
 XX  
 XX 09-APR-2003; 2003WO-US010879.  
 XX  
 XX 10-APR-2002; 2002US-00121258.  
 PR 29-JUL-2002; 2002US-00209208.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 PI Tsien RY, Campbell RE, Baird GS;  
 P-PSDB; ADL46203.  
 DR  
 DR MPI: 2003-845265/78.  
 XX  
 PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
 PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
 PT in molecular biology, e.g. in immunoassays or in tracking protein  
 PT movement in cells.  
 XX  
 PS Disclosure; SEQ ID NO 2; 166pp; English.  
 PS  
 XX The invention relates to a polynucleotide sequence encoding a Discosoma  
 CC red fluorescent protein (DRed) variant having a reduced propensity to  
 CC oligomerize. The protein variant comprises one or more amino acid  
 CC substitutions at the AB and/or AC interface(s) of the wild-type DRed  
 CC sequence, where the substitutions result in reduced propensity of the  
 CC DRed variant to form tetramers and where the variant displays detectable  
 CC fluorescence of at least one red wavelength. The composition and methods  
 CC are useful in producing red fluorescent proteins having reduced  
 CC propensity for oligomerization, especially tetramerization. The protein  
 CC may be used in molecular biology and in other scientific applications,  
 CC such as in immunoassays or hybridization assays, or in tracking the  
 CC movement of proteins in cells. This sequence corresponds to the DRed  
 CC coding sequence.  
 XX  
 SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.3e-140 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0  
 US-10-006-922a-12 (1-225) x ADL46204 (1-678)



```

QY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheValArgMetGluGly 20
XX 1 ATGAGGCTCTCCAGAAATGTTATCAAGAGCTTCATGAGGTTTAAGGTTGCAATGAAAGGA 60
DB 21 ThrValAsnGlyVHISGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40
QY 61 ACGGTCAATGGGCAACGATTGAAATGAAAGGCAAGAGGAGGCGCATACGAAGGCG 120
DB 41 HisAsnThrValIleLeuLeuValThrLySGlyGlyProLeuProPheAATPAApIle 60
QY 121 CACAAATACCGTAAAGCTTAAGGTACCAAGGGGAGCTTTGGCATTTGCTGGGATATT 180
DB 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValIleHisProA1aApIlePro 80
QY 181 TTGTACCAACAATTTCAATGATGGAACCAAGTATATGTCAAGACCCCTGCCGACATACCA 240
DB 81 AspTyrLyLeuLeuSerPheProGluGlyPheLyTyrGluArgValMetAsnPheGlu 100
QY 241 GACTATATAAAAGCTGCATTTCTCGAAGGATTTAAATGGGAAAGGGTCATGAACTTTGAA 300
DB 101 AAspGlyValValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
QY 301 GACGGTGGCGTCTGCTACTGTAACCCAGATTCCAGTTTGCAAGATGCGTCTTCATCTAC 360
DB 121 LyValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnLySerThr 140
QY 361 AAGGTCAAGTTCATTTGGCGTGAACCTTTCTCCGATGGAACCTGTTATGCAAAAGAAAGACA 420
DB 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
QY 421 ATGGCGTGGGAAGCCAGCACTGAGCGTTGTATCTCGTATGCGCGTGTGAAAGAGAG 480
DB 161 IleHisIleValIleLeuLeuLyAsnAspGlyValHisTyrLeuValGluPheLySerIle 180
QY 481 ATTCATTAAGGCTCTGAAGCTGAAGACGGTGTGCTATTAAGTTGAATTCAGAAAGTATT 540
DB 181 TyrMetAlaIleValIleProValGlnLeuProGlyTyrTyrTyrValAspSerLyLeuAsp 200
QY 541 TACATGGCAAAAGAGCTGTGAGCTACCAAGGTAATCTATGATCTCCAAATCGAAT 600
DB 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
QY 601 ATTAACAAGCCACAACAAGACTATACATGCTTGAGACATATGAAAGAACCGAGGAGCGC 660
QY 221 HisHisLeuPheLeu 225
DB 661 CACCATCTGTCTCTT 675

```

## RESULT 13

ADN33978

ADN33978 standard; DNA; 678 BP.

ADN33978;

01-JUL-2004 (first entry)

wild-type DERED encoding sequence.

Childarian; fluorescence resonance energy transfer; FRET; wild-type DERED;

dis.

DiscoSoma sp.

W02003054158-A2.

03-JUL-2003.

18-DEC-2002; 2002W0-US040539.

19-DEC-2001; 2001US-0341723P.

(UYCH-) UNIV CHICAGO.

XX Bevis B, Glick B;  
 FI WPI; 2003-569236/53.  
 XX P-PSDB; ADN33979.  
 DR Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent  
 PT mutant of a Childarian chromo- or fluorescent protein or its mutant,  
 PT useful for applications involving chromo- or fluorescent proteins.  
 PS Claim 8; SEQ ID NO 1; 65bp; English.

CC The present invention relates to nucleic acid that encodes a rapidly  
 CC maturing chromo- or fluorescent mutant of a Childarian chromo- or  
 CC fluorescent protein or its mutant. The protein is useful in applications  
 CC involving nucleic acid encoding a chromo- or fluorescent protein and is  
 CC useful for producing a chromo and/or fluorescent protein which involves  
 CC growing the cell, whereby the protein is expressed, and isolating the  
 CC protein substantially free of other proteins. The protein is useful in  
 CC applications involving chromo- or fluorescent protein and is useful as  
 CC PCR primers, hybridization probes, etc. The expression cassettes are  
 CC useful for synthesizing related proteins. The chromoproteins are useful  
 CC as coloring agents which are capable of imparting color or pigment to a  
 CC particular composition of matter e.g. food compositions, pharmaceuticals,  
 CC cosmetics, living organisms, e.g., animals and plants. The chromoproteins  
 CC may also find use as labels in analyte detection assays, e.g. assays for  
 CC biological analytes of interest and as selectable markers in recombinant  
 CC DNA applications, e.g. the production of transgenic cells and organisms.  
 CC The fluorescent proteins find use in a variety of different applications,  
 CC e.g. in fluorescence resonance energy transfer (FRET) applications, as  
 CC biosensors in prokaryotic and eukaryotic cells, in applications involving  
 CC the automated screening of arrays of cells expressing fluorescent  
 CC reporting groups by using microscopic imaging and electronic analysis, as  
 CC second messenger detectors, and in fluorescence activated cell sorting  
 CC applications and as in vivo marker in animals. The fluorescent proteins  
 CC also find use in protease cleavage assays. The proteins can also be used  
 CC in assays to determine the phospholipid composition in biological  
 CC membranes and as a fluorescent timer. The present sequence represents the  
 CC wild-type DERED encoding sequence.

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

## Alignment Scores:

Score:	2,3e-140	Length:	678
Percent Similarity:	1214.00	Matches:	225
Percent Local Similarity:	100.00%	Conservative:	0
Query Match:	100.00%	Mismatches:	0
DB:	11	Indels:	0
		Gaps:	0

US-10-006-922A-12 (1-225) x ADN33978 (1-678)

```

QY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheValArgMetGluGly 20
XX 1 ATGAGGCTCTCCAGAAATGTTATCAAGAGCTTCATGAGGTTTAAGGTTGCAATGAAAGGA 60
DB 21 ThrValAsnGlyVHISGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40
QY 61 ACGGTCAATGGGCAACGATTGAAATGAAAGGCAAGAGGAGGCGCATACGAAGGCG 120
DB 41 HisAsnThrValIleLeuLeuValThrLySGlyGlyProLeuProPheAATPAApIle 60
QY 121 CACAAATACCGTAAAGCTTAAGGTACCAAGGGGAGCTTTGGCATTTGCTGGGATATT 180
DB 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValIleHisProA1aApIlePro 80
QY 181 TTGTACCAACAATTTCAATGATGGAACCAAGTATATGTCAAGACCCCTGCCGACATACCA 240
DB 81 AspTyrLyLeuLeuSerPheProGluGlyPheLyTyrGluArgValMetAsnPheGlu 100
QY 241 GACTATATAAAAGCTGCATTTCTCGAAGGATTTAAATGGGAAAGGGTCATGAACTTTGAA 300
DB 101 AAspGlyValValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

```

```

Db      301 GACGGTGGCGTCGTACTGTAACCAAGATTCCAGTTTGCAAGATGGCTGTTTCACTAC 360
Qy      121 LysValIysPheIleGIYValIAsnPheProSerAspGIYProValMetGlnLysThr 140
Db      361 AAGGCAAGTTCATTGGCGGTGAACTTTCCTTCGATGACCTGTTATCCAAAGAGACA 420
Qy      141 MetGIYTPGluAlaSerThrGluArgLeuTYrProArgAspGIYValIleuLysGIYGu 160
Db      421 ATGGGCTGGAAAGCCAGCACTGAGCGTTTGATCCTCGTGAATGGCGTGTGAAGAGAG 480
Qy      161 ILehIysAlaLeuLysLeuLysAspGIYGIYhiSTYrLeuValGIuPheLysSerIle 180
Db      481 ATTCAATAAGGCTCGTAAGCTGAAGACGGTGGTCAATTCCTAGTTGAATTCAAAAGTATT 540
Qy      181 TyrMetAlaLysLysProValGIuLeuProGIYTYrTYrValAspSerLysLeuAsp 200
Db      541 TACATGGCAAGAACCTGTGTGACCTACAGGGTACTACTATGTGACTCCAAACTGGAT 600
Qy      201 ILehTrSerhiAsnGIuAspTYrThrIleValGIuGIuInTYrGIuArgThrGIuGIYArg 220
Db      601 ATAACAAGCCCAACGAAGACTATACATCGTTGAGCAGTATGAAGAACGAGGAGACC 660
Qy      221 HisHisLeuPheLeu 225
Db      661 CACCATCTGTTCCCTT 675

RESULT 14
AD136420
ID      AD136420 standard; DNA; 678 BP.
AC      AD136420;
DT      22-APR-2004 (first entry)
DE      Discosoma sp. red fluorescent protein (RED) DNA.
XX      Fluorobody; binding ligand; green fluorescent protein; GFP;
KW      target detection; gene; red fluorescent protein; ds; RED.
OS      Discosoma sp.
XX      Key      Location/Qualifiers
FH      CDS      1..678
FT      /tag= a
FT      /product= "Discosoma sp. red fluorescent protein (RED) "
XX      US2003203355-A1.
XX      30-OCT-2003.
XX      24-APR-2002; 2002US-00132067.
XX      24-APR-2002; 2002US-00132067.
XX      (LALA-) LOS ALAMOS NAT LAB.
XX      (RBGC ) UNIV CALIFORNIA.
XX      Bradbury AM, Zeytun A, Waldo GS;
XX      WPI; 2004-154325/15.
XX      DR      P-PSDB; AD136421.
XX      Novel binding ligand with intrinsic fluorescence and comprising
XX      fluorescent protein having heterologous binding sites, useful for
XX      detecting target molecule.
XX      Example 6; SEQ ID NO 3; 23bp; English.
XX      The invention relates to binding ligands (fluorobodies) with intrinsic
XX      fluorescence, which comprises green fluorescent protein (GFP) having
XX      heterologous binding sites. The binding ligand is useful for detecting
XX      the target molecule and its efficiently detects the target molecule. The
XX

```

```

CC      present sequence is Discosoma sp. red fluorescent protein (RED) DNA used
CC      in the exemplification of the invention.
XX      SQ      Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      2.3e-140      Length:      678
Score:      1214.00      Matches:      225
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      12      Gaps:      0

US-10-006-922A-12 (1-225) x AD136420 (1-678)
Qy      1 MetArgSerSerLysAsnValIleLysGIuPheMetArgPheLysValArgMetGIuGIY 20
Db      1 ATGAGGCTCTTCCAAAGATGTTATCCAGAGAGTTCATGAGGTTTAAGTTCCGATGGAAGA 60
Qy      21 ThrValAsnGIYHisGIuPheGIuIleGIuGIYGIuGIYArgProTYrGIuGIY 40
Db      61 ACGGTCATGGGCGACGAGTTTGAAATGAAGCGCAAGAGAGGAGGCCATACGAAGGC 120
Qy      41 HisAsnThrValIysLeuLysValIThrLysGIYGIYProLeuProPheIleTPAspIle 60
Db      121 CACAATACCGTAAAGCTTAAAGTTAACCAAGGGGAGACCTTGCCATTGGCTTGGGATATT 180
Qy      61 LeuSerProGlnPheGlnTYrGIYSerLysValIYrValIysHisPProIAspIlePro 80
Db      181 TTGTCAACCAATTTCCAGTATGGAAGCAAGGTATATGTCAAGCACTCGCGACATACCA 240
Qy      81 AspTYrLysLysLeuSerPheProGIuGIYPhelYrTPGluArgValMetAsnPheGIu 100
Db      241 GACTATATAAAGCTGTGATTTCCGAAAGATTTAATGGAAAGGTCATGAACTTTGAA 300
Qy      101 AspGIYGIYValValIThrValIThrGlnAspSerSerLeuGlnAspGIYCyPheIleTYr 120
Db      301 GACGGTGGCGTCGTACTGTAACCAAGATTCCAGTTTGAGAGATGGCTGTTTCATCTAC 360
Qy      121 LysValIysPheIleGIYValIAsnPheProSerAspGIYProValMetGlnLysLysThr 140
Db      361 AAGGTCAAAGTTCATTGGCGTGAACCTTCCGATGACCTGTATGCAAAAGATATT 420
Qy      141 MetGIYTPGluAlaSerThrGluArgLeuTYrProArgAspGIYValIleuLysGIYGu 160
Db      421 ATGGGCTGGAAAGCCAGCACTGAGCGTTTGATCCTGTGATGGCGTGTGAAGAGAG 480
Qy      161 ILehIysAlaLeuLysLeuLysAspGIYGIYhiSTYrLeuValGIuPheLysSerIle 180
Db      481 ATTCAATAAGGCTCGAAGCTGAAGAACGGTGGTCAATTCCTAGTTGAATTCAAAAGTATT 540
Qy      181 TyrMetAlaLysLysProValGIuLeuProGIYTYrTYrValAspSerLysLeuAsp 200
Db      541 TACATGGCAAGAACCTGTGTGACCTACAGGGTACTACTATGTGTACTCCAAACTGGAT 600
Qy      201 ILehTrSerhiAsnGIuAspTYrThrIleValGIuGIuInTYrGIuArgThrGIuGIYArg 220
Db      601 ATAACAAGCCCAACGAAGACTATACATCGTTGAGCAGTATGAAGAACGAGGAGACC 660
Qy      221 HisHisLeuPheLeu 225
Db      661 CACCATCTGTTCCCTT 675

RESULT 15
ADM97768
ID      ADM97768 standard; DNA; 678 BP.
AC      ADM97768;
DT      01-JUL-2004 (first entry)
DE      D sp red fluorescent protein coding sequence SEQ ID NO: 21.
XX

```

KW de; gene; enzyme; sensor cell; fluorescent protein;  
 KW signal transduction detection system; promoter; targeting sequence;  
 KW targeted drug.

OS Discosoma sp.

Key Location/Qualifiers  
 CDS 1..678  
 FT /tag= a  
 FT /product= "fluorescent protein"

W02004031415-A2.

15-APR-2004.

05-SEP-2003; 2003WO-US028076.

05-SEP-2002; 2002US-0408297P.

(VERT-) VERTEX PHARM INC.

Whitney MA, Zeh K, Sanders PS,

WPI; 2004-330208/30.

P-PSDB; ADM97769.

Developing a sensor cell, useful in determining the activity of a target gene and in developing therapeutic drugs, comprises providing cells comprising a signal transduction detection system and introducing DNA construct into cells.

Disclosure; Page 167-168; 234pp; English.

The present invention relates to a method of developing a sensor cell, for determining the activity of a target gene in the cell, which comprises providing a homogeneous population of cells, where each of the cells comprises a signal transduction detection system and introducing into the population of cells an isolated DNA construct comprising a promoter operatively linked to a targeting sequence. The method is useful in developing a sensor cell for determining the activity of a target gene in the cell. The sensor cell and the methods are useful in developing new and therapeutic drugs directed to the targets. The present sequence is a coding sequence shown in the exemplification of the invention.

Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,3e-140 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-10-006-922a-12 (1-225) x ADM97768 (1-678)

QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20  
 DB 1 ATGAGGCTCTCCAGAGATGTTATCAGAGATTTCATGAGTTTAAGTCCGATGGAAGA 60  
 QY 21 ThrValAsnGlyHsGluPheGluIleGluGlyGluGlyGluGlyArgProIyGluGly 40  
 DB 61 ACGGTCAATGGGCGAGGTTTGAATATGAAGCGCAAGAGAGAGGAGGCGATAGGAAGC 120  
 QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProIeuProPheIleATPAspIle 60  
 DB 121 CACAAATACCGTTAAAGCTTAAGTAACCAAGGGGGGACCTTGGCATTTGCTGGGATATT 180  
 QY 61 LeuSerProGlnPheGlnIyGlySerIysValIyValIyValIyValIyValIyValIy 80  
 DB 181 TTGTCAACCAATTTCAATGATGGAAGCAAGTATATGTCAAGCACTTCCGACATACCA 240  
 QY 81 AspTyrIleValIysLeuSerPheProGluGlyPheIyIleTyrGluArgValMetAsnIleGlu 100

DB 241 GACTATATAAAAGCTGCTATTTCTCGAAGATTTAATGGAAGGTCATACCTTGAA 300  
 QY 101 AspGlyValValThrValThrGlnAspSerSerIeuGlnAspGlyCysPheIleTyr 120  
 DB 301 GACGGTGGCGTGGTACTGTAACCAAGATTCCAGTTTGACAGATGGCTGTTGCATCTAC 360  
 QY 121 LysValIysPheIleGlyValAsnPheProSerIeuAspGlyProValMetGlnIyIysThr 140  
 DB 361 AAGTCAGTTCATTTGGGGTGAACCTTCTCCAGTGAACCTGTTATGCAAAAGAGACA 420  
 QY 141 MetGlyTyrPgluAlaSerThrGluArgLeuTyrProAspAspGlyValIleuIysGlu 160  
 DB 421 ATGGCTGGAAAGCCAGCACTGAGCGTTTGATCTCTGATGGCTGTGAAAGAGAG 480  
 QY 161 IleHisValAlaLeuIysLeuIysAspGlyIyHisTyrIleValGluPheIysSerIle 180  
 DB 481 ATTCATTAAGGCTGTGAAGCTGAAAGACGGTGGTCATTAAGTTGAATTCAAAAGTATT 540  
 QY 181 TyrMetAlaIyIysPheProValGlnIeuProGlyTyrTyrTyrValAspSerIyIysAsp 200  
 DB 541 TACATGGCAAAAGAACCTGTGCACTACAGGATCTATGTTGACTCCAAACTGGAT 600  
 QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnIyGluArgThrGluGlyArg 220  
 DB 601 ATACAGGCCACACAGAAAGACTATACATTCGTTGACGATGAAAGAACGAGGAGCGC 660  
 QY 221 HisHisLeuPheLeu 225  
 DB 661 CACCATCTGTTCTT 675

Search completed: January 12, 2006, 10:16:26  
 Job time : 485 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 12, 2006, 09:14:42 / Search time 3744 Seconds  
(without alignment)  
2811.724 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFRKFRKMEG.....EDYIVEQYERTEGRHFL 225

Scoring table:  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus.p2n.model -DEV=rlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10006922/runat\_10012006\_162408\_18933/app\_query.fasta\_1.391  
-DB=EST -QFMT=fasterp -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNIT9=bits -START=1 -END=1 -MATRIX=biolenum62 -TRANS=human40.cdi -LIST=45  
-OCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922.@CNCN 1 1 8010 @runat\_10012006\_162408\_18933 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEDUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_est3:\*
- 4: gb\_hic:\*
- 5: gb\_est4:\*
- 6: gb\_est5:\*
- 7: gb\_est6:\*
- 8: gb\_est7:\*
- 9: gb\_gse1:\*
- 10: gb\_gse2:\*
- 11: gb\_gse3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	461	38.0	556	DR987865 JGI_AOSF1
2	461	38.0	556	DR987866 JGI_AOSF1
3	225	18.5	661	B1385436 BFL26_000
4	218.5	18.0	508	BW802012 BW802012
5	218.5	18.0	646	BW736869 BW736869
6	217.5	17.9	639	BW736108 BW736108
7	216.5	17.8	604	BW738638 BW738638

Result No.	Score	Query Match Length	ID	Description
8	216.5	17.8	606	BW741633 BW741633
9	216.5	17.8	637	B1379468 BFLG1_001
10	216.5	17.8	687	B1387917 BFL26_002
11	212.5	17.5	643	B1386699 BFL26_001
12	212.5	17.5	675	B1382638 BFLG2_000
13	210.5	17.3	590	BW883955 BW883955
14	210.5	17.3	812	BW893733 BW893733
15	209.5	17.3	790	AL044652 DKFZP434P
16	206.5	17.0	630	BW731527 BW731527
17	206.5	17.0	677	BW894392 BW894392
18	205.5	16.9	605	BW744973 BW744973
19	205.5	16.9	618	BW871203 BW871203
20	205.5	16.9	668	BW894220 BW894220
21	203.5	16.8	555	BW845436 BW845436
22	203.5	16.8	602	BW872238 BW872238
23	203.5	16.8	620	BW870345 BW870345
24	203.5	16.8	630	BW740938 BW740938
25	203.5	16.8	641	BW798018 BW798018
26	203.5	16.8	641	BW881726 BW881726
27	203.5	16.8	647	BW736374 BW736374
28	203.5	16.8	652	BW733673 BW733673
29	203.5	16.8	652	BW881040 BW881040
30	202.5	16.7	602	BW870882 BW870882
31	202.5	16.7	744	DR622961 EST101308
32	201.5	16.6	573	DR622961 EST101308
33	201.5	16.6	675	BW865453 BW865453
34	200.5	16.5	722	DR624559 EST101358
35	200.5	16.5	730	DR623458 EST101358
36	200.5	16.5	736	DR629433 EST101943
37	200.5	16.5	768	DR629433 EST101943
38	200.5	16.5	791	DR621827 EST101195
39	200.5	16.5	795	DR624841 EST101496
40	200.5	16.5	818	DR627300 EST101742
41	200.5	16.5	828	DR627300 EST101742
42	200.5	16.5	831	DR628628 EST101875
43	200.5	16.5	833	DR624844 EST101497
44	200.5	16.5	840	DR624776 EST101490
45	200.5	16.5	862	DR624009 EST101413

## ALIGNMENTS

RESULT 1  
LOCUS DR987865/c 556 bp mRNA linear EST 03-AUG-2005  
DEFINITION JGI\_AOSF1131.rev AOSF Montastraea faveolata adult colony  
Montastraea faveolata CDNA clone AOSF1131 3', mRNA sequence.

ACCESSION DR987865.1 GI:71782571  
VERSION  
KEYWORDS  
SOURCE Montastraea faveolata  
ORGANISM Montastraea faveolata  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
Favina; Faviidae; Montastraea.

REFERENCE 1 (bases 1 to 556)  
Schwarz,J.A., Brokstein,P., Manohar,C., Coffroth,M.A., Szman,A.  
and Medina,M.  
TITLE Coral-Symbiodinium EST Project  
JOURNAL Unpublished (2005)  
COMMENT Other ESTs: JGI\_AOSF1131.fwd  
Contact: Schwarz, JA, Medina, M.  
Evolutionary Genomics  
DOE Joint Genome Institute  
2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
Tel: 925-296-5823  
Email: jascwarz@lbl.gov

CDNA Library Preparation: DOE Joint Genome Institute:  
http://www.jgi.doe.gov  
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov  
Naming Conventions: EST name is generated by the concatenation of  
the JGI Clone Id and the direction of sequencing. The suffix '.rev'  
indicates a reverse sequencing read of the insert. It does not  
necessarily reflect the orientation of the insert.





Db	599	AAGAGGACGCGGCTTCTGTGTGATCCGCAAGTGTACTGTAACGCTTCCAAAGCCGAATC	658
Qy	202	Thr 202	
Db	659	ACC 661	
RESULT 4			
BW802012			
LOCUS			
DEFINITION	BW802012	508 bp	mRNA
ACCESSION	BW802012		linear
VERSION	BW802012.1		EST 23-MAY-2005
KEYWORDS	larva whole animal	Branchiostoma floridae	unpublished cDNA library
SOURCE	mRNA sequence.	Branchiostoma floridae	CDNA clone bblv023d19 5',
ORGANISM	Branchiostoma floridae	(Florida lancelet)	
REFERENCE	Branchiostoma floridae		
AUTHORS	Y.,J., Holland,L.Z., Shin-I,T., Kohara,Y., Satou,Y. and Satoh,N.		
TITLE	Expressed genes in Branchiostoma floridae		
JOURNAL	Unpublished (2005)		
COMMENT	Contact: Tadasu Shin-i		
	Center for Genetic Resource Information		
	National Institute of Genetics		
	111 Yata, Mishima, Shizuoka 411-8540, Japan		
	Tel: 81-559-81-6856		
	Fax: 81-559-81-6855		
	Email: tshin@genes.nig.ac.jp.		
FEATURES	Location/Qualifiers		
source	1..508		
	/organism="Branchiostoma floridae"		
	/mol_type="mRNA"		
	/db_xref="taxon:7739"		
	/clone="bblv023d19"		
	/division_type="whole animal"		
	/dev_stage="larva"		
	/clone_idb="Amphioxus Branchiostoma floridae unpublished		
	CDNA library, larva whole animal"		
ALIGNMENT SCORES:			
Pred. No.:	8.98e-16	Length:	508
Score:	218.50	Matches:	52
Percent Similarity:	51.41%	Conservative:	21
Best local Similarity:	36.62%	Mismatches:	50
Query Match:	18.00%	Indels:	19
DB:	5	Gaps:	4
US-10-006-922A-12 (1-225) x BW802012 (1-508)			
Qy	18	MecGluglYThrValaenGIYhIeGIuPhelGluIleGIuGIYgluGIYalxPro	37
Db	103	ATGTTGGCTTCATCAACGCGCAAGAGTTCACCTGGGGCGGTGAAGAGGACCGG	162
Qy	38	TyrGIuGIYhIaenThrValIYleuGIYValThrlYsGIYglYProLeuProPhala	57
Db	163	AATGACGGAGCGCTGAGAGCAAGAGTGGCGTCACTAAGGAGGCC---CTGCCCTTCTCC	219
Qy	58	TrpAerIleLeuSerProGlnPhelGlnTyrGlySerIYValIYhIeProAla	77
Db	220	CCCGTGAATCCCGGCCCTTAACCTGGGAGTACCGGATACCAACGATACCTGCCCTCCGAGCC	279
Qy	78	AspIleProAerIYhIYleuLeuSerPheProGIuGIYhIeYstrGIuIangValMet	97
Db	280	GGAACCTCAACCGTACCGAGGACCATCAACGAGAGGATACCAAAAGCACCGACCTTC	339
Qy	98	AsnPhelGluAerGIYglYValThrlYValThrlGlnAerSerSerLeuGlnAerGIYCyS	117
Db	340	AAGTTCAAGAGCGCGGCGTATGACCATCAAC-----	372

```

Oy 118 Phaeletyr-----LysVallys-----PheIleGlyValalan 128
Db 373 TTCGGCTACACCTACTCAGGACCAAGATCAAGAGGAGTTCACGTGTGATGCCGG 432
Oy 129 PheProSerAapGlyProValMetGlnLysThrMetGlyTrpGluAlaSerThrglu 148
Db 433 TTCCTATATGACGGCCCTGTATATGACCAACTCATCTCAGCAGCAGATCACAACGTGGAG 492
Oy 149 ArgLeu 150
Db 493 AGGCTG 498

RESULT 5
BW736869
LOCUS
DEFINITION
Accession
Version
Keywords
Source
Organism
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
SOURCE
FEATURES
Source
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-10-006-922A-12 (1-225) x BW736869 (1-646)
Oy 16 ValArgMetGluGlyThrValaAnglyVHIGluBheGluIleGluGlyGluGly 35
Db 61 ATCCACCTTCACGGCTCCATCAACGGCCAGCAGATTCCGACATGTCGGGAGGAGAAAGGC 120
Oy 36 ArgProGlyTrpGlyVHIGluBenthrValLysLeuLysValThrLysGlyGlyProLeuPro 55
Db 121 GACCCGAAGCCGGCTCGCTGTCGACCAACAGCAAAATCCACCAAG--GGTCCCTGAAG 177
Oy 56 PheAlaIrrApeAlaIleLeuSerProGlnBheGlnIrrGlySerLysValIrrValLysHs 75
Db 178 TTCCTCTCCCACTGATGATATCCCCACCTCGGGGTACGGGTACTACCACTACTCTCCCTTAC 237

```



```

Qy 76 ProAlaApIleProAspTyrIleValLeuSerPheProGlu-----GlyPheLysTrp 93
Db 228 CCG---GACGGAACCTCGCTTTCACAGCCACCATGTTGAAGATCGGGGTATACAGTC 294
Qy 94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerLeu 113
Db 225 CACCGCGTGTTCGACTTGAAGACGAGGCAAGCTGCTCATCGAGTTCAAGTACTCTAC 354
Qy 114 GlnAspGlyCySPheIleTyrIleValLysPheIleGlyValAsnPheProSerAspGly 133
Db 355 GAGGCTTCCCATATTAAAGCCGACATGAAATTACGGAACCGGTTTCCCTGAGAACGG 414
Qy 134 ProValMetGlnLysPheThrMetGlyTyrGlnAlaSer---ThrGluArgLeuTyrPro 152
Db 415 CCGGTCATGACGACGACGATTTGTGACGACGAGCGGTGCTGTCAGAAACCTACCTG 474
Qy 153 ArgAspGlyValLeuLysGlyGlnIleHISylsAlaLeuLysLeuLysAspGlyGlyHis 172
Db 475 AACGACAACACCATGTGAGACAACTTCGACTGACCTTCAACCTTCAGAAATGGAAAGCGC 534
Qy 173 TyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186
Db 535 TACAGGCGCCGAGGTGAGCCATCACTTTCGACAAAGCCC 576

```

```

RESULT 6      639 bp  mRNA  linear  EST 09-AUG-2005
BM736108      Branchiostoma floridae unpublished cDNA library,
LOCUS         egg whole animal Branchiostoma floridae cDNA clone dbeg023022 5',
DEFINITION    mRNA sequence.
ACCESSION     BM736108.1 GI:66322738
VERSION       BM736108
KEYWORDS      EST.
SOURCE        Branchiostoma floridae (Florida lancelet)
ORGANISM      Branchiostoma floridae
              Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
              Branchiostoma.

```

```

REFERENCE     1 (bases 1 to 639)
AUTHORS       Yu,J., Holland,L.Z., Shin-I,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE         Expressed genes in Branchiostoma floridae
JOURNAL       Unpublished (2005)
COMMENT       Contact: Tadaeu Shin-I
              Center For Genetic Resource Information
              National Institute of Genetics
              111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshin@genes.nig.ac.jp
              If you want to have a cDNA clone for this EST or if you have any
              questions, please send an e-mail to Nori Satoh
              (satoh@acidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou
              (yutaka@acidian.zool.kyoto-u.ac.jp).
              Location/Qualifiers
                1..639
                  /organism="Branchiostoma floridae"
                  /mol_type="mRNA"
                  /db_xref="taxon:7739"
                  /clone="dbeg023022"
                  /issue_type="whole animal"
                  /dev_stage="egg"
                  /clone_lib="Amphioxus Branchiostoma floridae unpublished
                  cDNA library, egg whole animal"

```

```

FEATURES
  source
    /organism="Branchiostoma floridae"
    /mol_type="mRNA"
    /db_xref="taxon:7739"
    /clone="dbeg023022"
    /issue_type="whole animal"
    /dev_stage="egg"
    /clone_lib="Amphioxus Branchiostoma floridae unpublished
    cDNA library, egg whole animal"

```

# ORIGIN

```

Alignment Scores:
Pred. No.:      1,64e-15      Length:      639
Score:          217.50        Matches:      54
Percent Similarity: 48.85%    Conservative: 31
Best Local Similarity: 31.03%  Mismatches:    84
Query Match:    17.92%      Indels:        5
DB:              5          Gaps:          4

```

US-10-006-922a-12 (1-225) x BM736108 (1-639)

```

Qy 16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGly 35
Db 67 ATCAACCTCCACGCGCTCATGACGACGACGATTCGACCTGGTGGGGGAGAAAGGC 126
Qy 36 ArgProTyrGluGlyHisAsnThrValLysLeuValThrLysGlyGlyProLeuPro 55
Db 127 GACCGAAGCGCGCTGCTGTGACCAACAGCAAAATCCACCAAG---GGTCCCTGAAG 183
Qy 56 PheIleTyrAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHis 75
Db 184 TTCTCTCCCTTCACTGATGATATCCCACTCGGAGTACGGGTACTACAGTACTCTCCCTAC 243
Qy 76 ProAlaApIleProAspTyrIleValLeuSerPheProGlu-----GlyPheLysTrp 93
Db 244 CCG---GACGGAACCTCGCTTTCAGGCGCACCATGTTGGAAGAAACAGGGGTATGCAAGTC 300
Qy 94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerLeu 113
Db 301 CACCGCGTGTTCGACTTGAAGACGAGGCAAGCTGTCATGAGACTTTAAGTACTCTAC 360
Qy 114 GlnAspGlyCySPheIleTyrIleValLysPheIleGlyValAsnPheProSerAspGly 133
Db 361 GAGGTTCCCATATTAAAGCCGACATGAAATTACAGGGAACCGGTTTCCCTGAGAACGG 420
Qy 134 ProValMetGlnLysPheThrMetGlyTyrGlnAlaSer---ThrGluArgLeuTyrPro 152
Db 421 CCGGTCATGACGACGACGATTTATGACGACGAGCGGTGCTGTCAGAAACCTACCTT 480
Qy 153 ArgAspGlyValLeuLysGlyGlnIleHISylsAlaLeuLysLeuLysAspGlyGlyHis 172
Db 481 AACGACAACACCATGTGAGACGCTTCGACTGACCTTCAACCTTCAGAAATGGAAAGCGC 540
Qy 173 TyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186
Db 541 TACAAAGCCCAAGTACGACGACCATCACTTTCGCGCAAGCCC 582

```

```

RESULT 7      604 bp  mRNA  linear  EST 09-AUG-2005
BM738638      Branchiostoma floridae unpublished cDNA library,
LOCUS         egg whole animal Branchiostoma floridae cDNA clone dbeg008a05 5',
DEFINITION    mRNA sequence.
ACCESSION     BM738638
VERSION       BM738638
KEYWORDS      EST.
SOURCE        Branchiostoma floridae (Florida lancelet)
ORGANISM      Branchiostoma floridae
              Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
              Branchiostoma.

```

```

REFERENCE     1 (bases 1 to 604)
AUTHORS       Yu,J., Holland,L.Z., Shin-I,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE         Expressed genes in Branchiostoma floridae
JOURNAL       Unpublished (2005)
COMMENT       Contact: Tadaeu Shin-I
              Center For Genetic Resource Information
              National Institute of Genetics
              111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshin@genes.nig.ac.jp
              If you want to have a cDNA clone for this EST or if you have any
              questions, please send an e-mail to Nori Satoh
              (satoh@acidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou
              (yutaka@acidian.zool.kyoto-u.ac.jp).
              Location/Qualifiers
                1..604
                  /organism="Branchiostoma floridae"
                  /mol_type="mRNA"
                  /db_xref="taxon:7739"
                  /clone="dbeg008a05"
                  /issue_type="whole animal"

```

# FEATURES

```

  source
    /organism="Branchiostoma floridae"
    /mol_type="mRNA"
    /db_xref="taxon:7739"
    /clone="dbeg008a05"
    /issue_type="whole animal"

```



SOURCE Branchiostoma floridae (Florida lancelet)  
 ORGANISM Branchiostoma floridae  
 Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
 Branchiostoma.  
 REFERENCE 1 (bases 1 to 637)  
 Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A., J.,  
 Herwig, R., Vingron, M. and Lehrach, H.  
 New evidence for genome-wide duplications at the origin of  
 vertebrates using an amphioxus gene set and completed animal  
 genomes  
 Genome Res. 13 (6A), 1056-1066 (2003)  
 JOURNAL PUBMED  
 12799346  
 COMMENT Contact: Panopoulou G  
 Laboratory 145, dept. Lehrach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihmestr.63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1235  
 Fax: +49 30 8413 1128  
 Email: panopoul@molgen.mpg.de  
 The library was characterised by oligonucleotide fingerprinting  
 (ONFP) to reduce sequencing redundancy. According to the ONFP  
 procedure, clones giving the same hybridisation pattern with a  
 battery of 200 oligonucleotides are grouped into clusters. One  
 clone per cluster is selected for sequencing. The size of each  
 cluster is an indicator of the frequency of a transcript in the  
 analysed library. The cluster size as well the coordinates of the  
 rest of the clones assigned to the same fingerprint cluster as the  
 clone from which the above EST is generated is available at the  
 amphioxus project site at: http://www.molgen.mpg.de/amphioxus/  
 Clones and filters are distributed via the Resource Center/Primary  
 Database of the German Genome Project (http://www.rzpd.de)  
 PCR Primers  
 FORWARD: 5' CCCAGGCTTACATTATGCTCCGGCTG 3' (M13RSP)  
 BACKWARD: 5' GCTATTACGCCAGCTCGGAGGGGATGTG 3' (M13RSP)  
 Insert Length: 1200 Std Error: 0.00  
 Seq primer: 5'-CCGTCGGAAATTCGCGGT-3' pSPORT3/86  
 High quality sequence stop: 637.  
 Location/Qualifiers  
 1..637  
 /organism="Branchiostoma floridae"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7739"  
 /clone="MPMP49801918"  
 /tissue\_type="whole embryo"  
 /dev\_stage="5-6 hrs (gastrula stage)"  
 /lab\_host="E.coli, XL1 blue"  
 /clone\_lib="Amphioxus 5-6 hrs cDNA library (Name  
 convention: BFL6 or MPMP498)"  
 /note="Vector: pSPORT1; Site 1: SalI, KpnI, EcoRI (5');  
 Site 2: NotI, BamHI, HindIII (3'); OligodT primed and  
 directionally cloned in pSPORT1 vector using a NotI  
 (5'-GGACTGATTCCTGATCGGAGCGCGGCC (T)15-3' and a SalI 5'-  
 TCGACCCAGCGCTCG-3' adapters (Gibco BRL)."  
 ORIGIN  
 Alignment Scores: Length: 637  
 Score: 216.50 Matches: 53  
 Percent Similarity: 48.57% Conservative: 32  
 Best Local Similarity: 30.29% Mismatches: 85  
 Query Match: 17.83% Indels: 5  
 DB: 3 Gaps: 4  
 US-10-006-922a-12 (1-225) x B1379468 (1-637)  
 Oy 15 LysValArgMetGluGlyThrValaAmGlyHisGluPheGluIleGluGlyGluGlyGlu 34  
 Db 79 GAGATCCACCTCCACGCTTCATCAACGCGCACAGTTCGAGCTGCTGCTGAGAAA 138  
 Oy 35 GlyArgProGlyTrpGluGlyHisAsnThrValLeuLeuValThrIleArgGlyGlyProLeu 54  
 Db 139 GGGACCCAAACCGCGCTCGTGTGTGACCAAGCGAAATCCACCCAG--GGTCCCTG 195

Oy 55 ProPheAlaTrpAspIleLeuSerProGlnPheGlnIleGlySerIleValIleValIys 74  
 Db 196 AAGTCTCTCCCACTGATGATCCCACTCGGGTACGGGTACTACGATCTCTCCC 255  
 Oy 75 HisProAlaAspIleProAspTyrIleValLeuSerPheProGlu-----GlyPheIys 92  
 Db 256 TACCCG---GACGAGCCCTCGCTTCCAGGCGCACATGTGGAAGATCGGGTATACA 312  
 Oy 93 TrpGluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSer 112  
 Db 313 GTCCACGCGGTGTTGACTTCGAAAGACGAGCGAAGCTTCACAGTTCACTACTCC 372  
 Oy 113 LeuGlnAspGlyCysPheIleTyrIleValIleValPheIleGlyValaAsnPheProSerAsp 132  
 Db 373 TACAGAGGTTCCCATATTAAAGCCGACATGAAAGTTACAGGAAACCGGTTCCCTGAGAT 432  
 Oy 133 GlyProValMetGlnIleValIleValIleValIleValIleValIleValIleValIleVal 151  
 Db 433 GGGCCGCTCATGACACGACCAATTTGTGACAGAGCGCTCGTGTCCAGAACACTAC 492  
 Oy 152 ProArgAspGlyValLeuIleValGlyGluIleHisIleValaLeuIleValLeuAspGlyGly 171  
 Db 493 CTCACAGCAACACACCATCGTGACAACTTCGACTGCACTTCAACCTGCAAGATGGGAAG 552  
 Oy 172 HisTyrLeuValGluPheIleValIleValIleValIleValIleValIleValIleValIle 186  
 Db 553 CGCTACAGAGGCCGAGTGACAGCACTTCCAGCAAGCCC 557  
 RESULT 10  
 LOCUS B1387917  
 DEFINITION BFL26\_002620 Amphioxus 26hr cDNA library (Name convention: BFL26 or  
 MPMP531) Branchiostoma floridae cDNA clone MPMP531IN0767 5', mRNA  
 sequence.  
 ACCESSION B1387917  
 VERSION B1387917.1 GI:30922753  
 KEYWORDS EST.  
 SOURCE Branchiostoma floridae (Florida lancelet)  
 ORGANISM Branchiostoma floridae  
 Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
 Branchiostoma.  
 REFERENCE 1 (bases 1 to 687)  
 Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A., J.,  
 Herwig, R., Vingron, M. and Lehrach, H.  
 New evidence for genome-wide duplications at the origin of  
 vertebrates using an amphioxus gene set and completed animal  
 genomes  
 Genome Res. 13 (6A), 1056-1066 (2003)  
 JOURNAL PUBMED  
 12799346  
 COMMENT Contact: Panopoulou G  
 Laboratory 145, dept. Lehrach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihmestr.63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1235  
 Fax: +49 30 8413 1128  
 Email: panopoul@molgen.mpg.de  
 The library was characterised by oligonucleotide fingerprinting  
 (ONFP) to reduce sequencing redundancy. According to the ONFP  
 procedure, clones giving the same hybridisation pattern with a  
 battery of 200 oligonucleotides are grouped into clusters. One  
 clone per cluster is selected for sequencing. The size of each  
 cluster is an indicator of the frequency of a transcript in the  
 analysed library. The cluster size as well the coordinates of the  
 rest of the clones assigned to the same fingerprint cluster as the  
 clone from which the above EST is generated is available at the  
 amphioxus project site at: http://www.molgen.mpg.de/amphioxus/  
 Clones and filters are distributed via the Resource Center/Primary  
 Database of the German Genome Project (http://www.rzpd.de)  
 PCR Primers  
 FORWARD: 5' CCCAGGCTTACATTATGCTCCGGCTG 3' (M13RSP)  
 BACKWARD: 5' GCTATTACGCCAGCTCGGAGGGGATGTG 3' (M13RSP)  
 Insert Length: 1 Std Error: 0.00  
 Seq primer: 5'-CCGTCGGAAATTCGCGGT-3' pSPORT3/86

High quality sequence stop: 687.

FEATURES  
Location/Qualifiers

1..687

/organism="Branchiostoma floridae"

/mol\_type="mRNA"

/db\_xref="taxon:7739"

/clone="MPMG531N0767"

/tissue\_type="whole embryo"

/dev\_stage="26 hrs (neurula stage)"

/lab\_host="E.coli, XLI blue"

/clone\_id="Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGP531)"

/note="Vector: pSport1 (Gibco BRL); Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-GACTAGTCTTAGATCGGAGCGCGGCC (T)15-3' and a SalI 5'-TCGACCCACGCGTCCG-3' adapters (Gibco BRL)."

Alignment Scores:  
Pred. No.: 2.41e-15 Length: 687  
Score: 216.50 Matches: 54  
Percent Similarity: 48.85% Conservative: 31  
Best Local Similarity: 31.03% Mismatches: 84  
Query Match: 17.83% Indels: 5  
Gaps: 4

ORIGIN

US-10-006-922a-12 (1-225) x B1387917 (1-687)  
16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGly 35  
::: ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
99 ATCCACCTCCACGGCTCCATCAAGCGCCACGAGTTGCACATGTCGGCGGAGGAAAGGC 158  
DB 36 ArgProGlyGluGlyHisAsnThrValIleLeuValIleThrIleGlyProLeuPro 55  
159 GACCCGAACCGCGCTGCTGTCGACCAACGAGCAATCCACCAAG--GGTCCCTGAAAG 215  
QY 56 PheIleThrPheIleLeuSerProGlnPheGlnIleGlySerIleValIleValIleHis 75  
DB 216 TTCTCTCCCTTACTGATGATCCCACTCGGGTACGGGTTCTACCACTACTCTCCCTAC 275  
QY 76 ProIleAspIleProAspIleTyrIleIleLeuSerPheProGlu-----GlyPheIleTyr 93  
DB 276 CCG---GACGGAACCTCGCTCCCTTCCAGGCCACATGTTGGAAGAACAGGATATGCAATC 332  
QY 94 GluArgValMetAsnPheGluAspGlyGlyValValIleThrValIleThrGlnAspSerIleu 113  
DB 333 CATGCGCTGTCGACTTTGAAGACGAGGCAAGCTGTCATCGAGTTTAAGTACTCTCTAC 392  
QY 114 GlnAspGlyCysPheIleTyrIleValIlePheIleGlyValAsnPheProSerAspGly 133  
DB 393 GAGGGTTCCCATATTCAGAGCCGACATGAAATTCACGGGAAACCGGTTCCCTGAGGACGGG 452  
QY 134 ProValMetGlnIleIleTyrThrMetGlyTyrGluIleAsp---ThrGluArgLeuTyrPro 152  
DB 453 CCGGTCATGACCAAGCAGATTGTGACCAAGAGCGCTGCTCAAGAAACACCTACTCTT 512  
QY 153 ArgAspGlyValIleLeuIleGlyGluIleHisIleValIleLeuIleValIleAspGlyGlyHis 172  
DB 513 AACCAACAACCATCTGTCAGACGCTTCGACTGCACTTCAACATTCGCAAGATGGGAAGCGC 572  
QY 173 TyrLeuValGluPheIleSerIleTyrMetAlaIleValIlePro 186  
DB 573 TACAGGCCCAAGTAGACAGCCACTTACTTCCGCAAGGCC 614RESULT 11  
LOCUS B1386699 643 bp mRNA linear EST 26-AUG-2003  
DEFINITION BFL26\_001402 Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGP531) Branchiostoma floridae cDNA clone MPMGP531J1370 5', mRNA sequence.  
ACCESSION B1386699  
VERSION B1386699.1 GI:30921604KEYWORDS  
SOURCE

ORGANISM

Branchiostoma floridae (Florida lancelet)  
Branchiostoma floridae  
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
Branchiostoma.REFERENCE  
AUTHORS 1 (bases 1 to 643)Panopoulou,G., Hennig,S., Groth,D., Krause,A., Pousetka,A.J.,  
Herwig,R., Vingron,M., and Lehrach,H.New evidence for genome-wide duplications at the origin of  
vertebrates using an amphioxus gene set and completed animal  
genome

Genome Res. 13 (6A), 1056-1066 (2003)

Genome Res. 13 (6A), 1056-1066 (2003)

Contact: Panopoulou G  
12793346Laboratory 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Innestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128

Email: panopoul@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting  
(ONPF) to reduce sequencing redundancy. According to the ONPF  
procedure, clones giving the same hybridisation pattern with a  
battery of 200 8mer oligonucleotides are grouped into clusters. One  
clone per cluster is selected for sequencing. The size of each  
cluster is an indicator of the frequency of a transcript in the  
analysed library. The cluster size as well the coordinates of the  
rest of the clones assigned to the same fingerprint cluster as the  
clone from which the above EST is generated is available at the  
amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>  
Clones and filters are distributed via the Resource Center/Primary  
Database of the German Genome Project (<http://www.rzpd.de>)PCR Primers  
FORWARD: 5' CCCGAGCTTTACACTTATGCTCCGCTCG 3' (M13RSP)  
BACKWARD: 5' GCTATTACGCGACGCTGCGCAAGGGGATGTG 3' (M13FSP)  
Insert Length: 1 Std Error: 0.00  
Seq primer: 5'-CCGGTCCGGAATTCGGCGT-3' pSport3/86

High quality sequence stop: 643.

Location/Qualifiers

1..643

/organism="Branchiostoma floridae"

/mol\_type="mRNA"

/db\_xref="taxon:7739"

/clone="MPMG531J1370"

/tissue\_type="whole embryo"

/dev\_stage="26 hrs (neurula stage)"

/lab\_host="E.coli, XLI blue"

/clone\_id="Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGP531)"

/note="Vector: pSport1 (Gibco BRL); Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-GACTAGTCTTAGATCGGAGCGCGGCC (T)15-3' and a SalI 5'-TCGACCCACGCGTCCG-3' adapters (Gibco BRL)."

ORIGIN

Alignment Scores:

Pred. No.: 6.69e-15 Length: 643

Score: 212.50 Matches: 56

Percent Similarity: 47.87% Conservative: 34

Best Local Similarity: 29.79% Mismatches: 87

Query Match: 17.50% Indels: 11

DB: 3 Gaps: 5

US-10-006-922a-12 (1-225) x B1386699 (1-643)

QY 16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGly 35  
DB 81 ATCCACCTCCACGGCTCCATCAAGCGCCACGAGTTGCACATGTCGGCGGAGGAAAGGC 140QY 36 ArgProGlyGluGlyHisAsnThrValIleLeuValIleThrIleGlyProLeuPro 55  
DB 141 GACCCGAACCGCGCTGTCGACCAAGCAATCCACCAAG--GGTCCCTGAAAG 197

Oy		56 PheAlaIeTrpApriLeUsErProGInPhEglnITrYcISerLySValTYrvAllyHis	75
Db		198 TTCTCCTTCAATGGATGAATGCCACCCTCGGTACGGGTACTACAAGTACTCCCTCAC	257
Oy		76 ProLlaAsPrlEPromAByTyLYSLysLeUSeRPHePrOglU-----GLYPheLySTrp	93
Db	Bt382638	258 CCG---GACGGAACCTTCGCCTTTCCAGGCCCTCCATTGTAAGAAGATCGGGGTATGCAGT	314
Oy		94 GluaArgVaIMecAnPheGlUAbePgLYGVlYaVAlThrVAItHrgInAbSerSerleU	113
Db		315 TACCGCGTGTCGACTTCCAAGAGCGAGGCGAACTACCGAGTTMAACTACTCCTAAC	374
Oy		114 GlnABepLYCySPheLIETryLValIVallYpHeLIegLYVaLaanPeProSeRASPdLY	133
Db		375 GAGGGTTCCCATATCAAAGCCGACATGAAGCTGATGGAGAGCGTTTTCCCTGACAGCGC	434
Oy		134 ProVALmeGlnULySYThMeLGLTyrPlUALasER---THrgUARgeUYtyFeRo	152
Db		435 CCAGTCAATGACCAAGCATGTTGTGCAGCAGAGCGGCTCGTGCAAGAAAGCATATCTT	494
Oy		153 ArgAPdeLYvAllEuLYSGLYGIUlILIsLYbaLAaleuLYSteuLYASpgLYGlyHS	172
Db		495 AACCAACAACAACATCTGACAGCTTCGACATGAGTTACAACCTGCATAATGGAAGCGC	554
Oy		173 TYrtreuvALGuIpHeLySErieLTymeALalySyLSvsPro-----	186
Db		555 TAGAGGGCCCCAGTGTCAAGCCACTACATCTTCGACAAAGCCCTTTTACAGCCGATCTCAT	614
Oy		187 ValGINteUProGilyTyTYrTYrTyr	194
Db		615 ATAAAGCAGCGCGGTCTTCGTGTAC	638
RESULT 12			
Bt382638			
LOCUS		BFHG2_000815 Amphioxus s-6 hrs cDNA library (Name convention: BFLG or MPKGps498) Branchiostoma floridae cdna clone MPKGP498F0258 5'	
DEFINITION		mRNA sequence.	
VERSION	Bt382638		
KEYWORDS	Bt382638.1 GI:30917704		
SOURCE	Eukaryote; Metazoa; Chordata; Cephalochordata; Branchiostomidae;		
ORGANISM	Branchiostoma floridae (Florida lancelet)		
	Branchiostoma floridae		
	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;		
	Branchiostoma		
	I (bases 1 to 675)		
REFERENCE	Papanicolaou,G., Hennig,S., Groch,D., Krause,A., Pousetka,A.J., Herwig,R., Vingron,M. and Lehnach,H.		
AUTHORS	New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes		
TITLE	Genome Res. 13 (6A), 1056-1066 (2003)		
JOURNAL	Contact: Panopoulou G		
PUBMED	Laboratory 145, dept.Lehnach		
COMMENT	Max-Planck-Institut fuer Molekulare Genetik Imnestr.63-73, D-14195 Berlin, Germany Tel.: +49 30 8413 1235 Fax: +49 30 8413 1128 Email: panopoul@molgen.mpg.de The library was characterized by oligonucleotide fingerprinting (ONPF) to reduce sequencing redundancy. According to the ONPF procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphiXus project site at: http://www.molgen.mpg.de/amphioxus/ Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (http://www.tzpd.de)		

```

FOR Primers
FORWARD: 5' , CCCCAGGCTTACATTANGCTTCGGGCTCG 3' (M13RSP)
BACKWARD: 5' , GCTATTACCCAGACGCTGCGAAGAAGGGGAGATGTG 3' (M13FSP)
Insert length: 1 Std Error: 0.00
Seq primer: 5' , -CCGGTCCGGAATTCGGGGGT-3' pSPORT3/86
High quality sequence stop: 675.
Location/Qualifiers

```

**FEATURES**  
**Source**

```

/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="MPMGp498F0258"
/tissue_type="whole embryo"
/dev_stage="5-6 hrs (gastrula stage)"
/lab_host="E.coli, XL1 blue"
/clone_lib="Amphioxus 5-6 hrs cDNA library (Name
convention: BRLG or MPMGP498)"
/note="Vector: pSPORT1; Site_1: SalI, KpnI, EcoRI (5'),
Site 2: NotI, BamHI, HindIII (3'); OligoDT primed and
directionally cloned in pSPORT vector using a NotI
(5'-pGACTAGTCTAGATGCGAGCGGCGGCC (r)15-3' and a SalI 5'-
TCGATCCACGGGTCG-3' adapters (Gibco BRL)."

```

## ORIGIN

Alignment Scores:		
Pred. No.:	7.17e-15	Length: 675
Score:	212.50	Matches: 54
Percent Similarity:	49.43%	Conservative: 32
Best Local Similarity:	31.03%	Mismatches: 83
Query Match:	17.50%	Indels: 5
DB:	3	Gaps: 4

US-10-006-922A-12 (1-225) x B1382638 (1-675)

OY	16	ValArgmetGIuGIYThrValAlaNGlyYHsGIbPhelGluIleGIuGIYGIuGIY	35
	110	ATTCACCTTCACGGCTCCATTCACAGCGCCAGAGTTGCACATGGTGGGGAGGAAAAAGCC	169
OY	36	ArgProTyrGIuGIYHsAsenThrValIleuYsValThryGIYGIYProIeuPro	55
Db	170	GACCCGMAAGCGGGCTCGCTGGTGCACACAGCGAAATCCACCAAG---GGTCCCTGGAAG	226
OY	56	PheAlaTrpApIleIeuSerProGlnPheGlnTyrGlySerYsValTyrValIleHis	75
Db	227	TTCTCCCTCACTGATGATCCCCACCTCGGGTACGGGTACTACCACTACCTCCCTAC	286
OY	76	ProAlaApIleProApIlyrTyrIlyLeuSerPheProGlu-----GlyPheTyr	93
Db	287	CCG---GACGGACCTCGCTTCCTTCCAGAGTCTCCATGTTGAAGATCGGGGTATGCAGTC	343
OY	94	GIuArgValIleAsnPhelGIuAspGIYGIYValValThrValThrGlnAspSerIeu	113
Db	344	TACCGCGTTCGACTTTGAAAGAGGAGGACCAAGCTACCGAGTTTAAAGTACTCTTAC	403
OY	114	GlnAspGIYCySphelIeTyrIlyValIlyPheIleGIYIleAlaPheProSerAspGIY	133
Db	404	GAGGGTTCCTCATTCAGAGCCGACATGAGCTAGTGGAGACGGTTCCTTCAGCAGCGC	463
OY	134	ProValIleGIuIlySlyrTherMetGIYTrpGluIleaser---ThrGluArgLeuTyrPro	152
Db	464	CGAGTCATGACCGACCGACATTTGTGCACCGAGCGCGCTCGCTCCMAAGAACGTATTT	523
OY	153	ArgAspGIYValIleuYsGIYGIuIleHisIlyAlaIleuYsIleuYsAspGIYGIYHs	172
Db	524	AACAAACAACACCATCGTGCAGAGCTTCGACTGCAGATTACAACCTGCAGAAATGGGAAAGCGC	583
OY	173	TyrIleuValGIuIlyPheYsSerIleTyrMetAlaIlySlyrPro	186
Db	584	TACAAAGGCCGAGTGTCAAGCCACTACATCTTCCAGCAAGCCC	625

RESULT 13

BW883955

LOCUS

590 bp

mRNA

linear

EST 24-MAY-2005

```

DEFINITION  BM83955 Amphioxus Branchiostoma floridae unpublished cDNA library,
neurula whole animal Branchiostoma floridae cDNA clone dbne135j09
ACCESSION   BM83955
VERSION     BM83955.1 GI:66498632
KEYWORDS    EST.
SOURCE      Branchiostoma floridae (Florida lancelet)
ORGANISM    Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE   1 (bases 1 to 590)
AUTHORS    Yu,J., Holland,L.Z., Shin-I,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE      Expressed genes in Branchiostoma floridae
JOURNAL    Unpublished (2005)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@gene.nig.ac.jp.
            Location/Qualifiers
FEATURES
    source          1..590
                     /organism="Branchiostoma floridae"
                     /mol_type="mRNA"
                     /db_xref="taxon:7739"
                     /clone="dbne135j09"
                     /tissue_type="whole animal"
                     /dev_stage="neurula"
                     /clone_lib="Amphioxus Branchiostoma floridae unpublished
                     cDNA library, neurula whole animal"
ORIGIN
Alignment Scores:
Pred. No.:      1.04e-14      Length:      590
Score:          210.50      Matches:      54
Percent Similarity: 48.85%      Conservative: 31
Best Local Similarity: 31.03%      Mismatches: 84
Query Match:    17.34%      Indels:      5
DB:             Gaps:      4

US-10-006-922a-12 (1-225) x BM83955 (1-590)
QY      16 ValArgMetGluGlyThrValaAnglyHfGluPheGluIleGluGlyGluGly 35
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      70 ATCCACCTTCACGGCTGCATCAACGCGCACGAGTTCGACATGGGGGAGAAAGGC 129
QY      36 ArgProTyrgluGlyHfAsnThrVallyLeuLyValThryGlyGlyProLeuPro 55
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      130 GACCGGAACGGCGGCTGCTGTGACCAACGAAATCCAAAG---GATCCCTGAAG 186
QY      56 PheAlaTrpAaplleLeuSerProGlnPheGlnTyrglySerLyValTyVallyeHis 75
       ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      187 TTCCTCCCTACTGATGATGCCCACTCGGGTACGCGGTACTACCACTACCTCCCTAC 246
QY      76 ProAlaAaplleProAapTyTrLyLyLeuSerPheProGlu-----GlyPheLyStrp 93
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      247 CCG---GACGGAACCTCGCTTCACAGGCTTCACATGTTGGAAGATCGGGGTATGCAATC 303
QY      94 GluArgValMetAsnPheGluAapGlyGlyValValThryValThrglnAapSerSerLeu 113
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      304 TACGCGGTGTCGACTTTGAAGACGAGGCAAGCTGACTACCGAGTTTAAGTACTCTCTAC 363
QY      114 GluAapGlyCySpheIleTyTrLyVallyAapheIleGlyValaAapheProSerAapGly 133
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      364 GAGGGTTCCTCATCAAGCGCAATCAAGGCTGAGGAAACGGGTTCCCTGACGACGGC 423
QY      134 ProAlaMetGlnLyLyPheThrMetClyTrPrGluAaSer---ThrgluArgLeuTyTrPro 152
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      424 CCAATCATGGCCACGACGATTTGTGACCAAGACGGCTGCTGTCACAAAGAACGTAATCTT 483
QY      153 ArgAapGlyValLeuLyGlyGluIleHfLyValaLeuLyLeuLyAapGlyGlyHis 172
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```

```

Db      484 AACAAACAACCATCTGTGACAGCTTCGACTGAGATTACAACTGCAGAAATGGAGCGC 543
QY      173 TyrlEuValGluPheLySerIleTyTrMetAlaLyLyPro 186
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      544 TACAGGGCCCGAGTGTGACGCACTACATCTTGCACAAAGCC 585

RESULT 14
BM893733
LOCUS     BM893733
DEFINITION  812 bp  mRNA  linear  EST 24-MAY-2005
            BM893733 Amphioxus Branchiostoma floridae unpublished cDNA library,
            neurula whole animal Branchiostoma floridae cDNA clone dbne123k07
            5', mRNA sequence.
ACCESSION   BM893733
VERSION     BM893733.1 GI:66511577
KEYWORDS    EST.
SOURCE      Branchiostoma floridae (Florida lancelet)
ORGANISM    Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE   1 (bases 1 to 812)
AUTHORS    Yu,J., Holland,L.Z., Shin-I,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE      Expressed genes in Branchiostoma floridae
JOURNAL    Unpublished (2005)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@gene.nig.ac.jp.
            Location/Qualifiers
FEATURES
    source          1..812
                     /organism="Branchiostoma floridae"
                     /mol_type="mRNA"
                     /db_xref="taxon:7739"
                     /clone="dbne123k07"
                     /tissue_type="whole animal"
                     /dev_stage="neurula"
                     /clone_lib="Amphioxus Branchiostoma floridae unpublished
                     cDNA library, neurula whole animal"
ORIGIN
Alignment Scores:
Pred. No.:      1.63e-14      Length:      812
Score:          210.50      Matches:      60
Percent Similarity: 46.92%      Conservative: 39
Best Local Similarity: 28.44%      Mismatches: 91
Query Match:    17.34%      Indels:      21
DB:             Gaps:      7

US-10-006-922a-12 (1-225) x BM893733 (1-812)
QY      16 ValArgMetGluGlyThrValaAnglyHfGluPheGluIleGluGlyGluGly 35
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      60 ATCCACCTTCACGGCTGCATCAACGCGCACGAGTTCGACATGGGGGAGAAAGGC 119
QY      36 ArgProTyrgluGlyHfAsnThrVallyLeuLyValThryGlyGlyProLeuPro 55
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      120 GACCGGAACGGCGGCTGCTGTGACCAACGAAATCCAAAGGGTGCC---CTGAAG 176
QY      56 PheAlaTrpAaplleLeuSerProGlnPheGlnTyrglySerLyValTyVallyeHis 75
       ||| :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      177 TTCCTCCCTACTGATGATGCCCACTCGGGTACGCGGTACTACCACTACCTCCCTAC 236
QY      76 ProAlaAaplleProAapTyTrLyLyLeuSerPheProGlu-----GlyPheLyStrp 93
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      237 CCG---GACGGAACCTCGCTTCACAGCTTCACATGTTGGAAGATCGGGGTATGCAATC 293
QY      94 GluArgValMetAsnPheGluAapGlyGlyValValThryValThrglnAapSerSerLeu 113
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      294 TACCGCGTGTTCGACTTTGAAGACGAGGCAAGCTGACTCCGAGTTTAAGTACTCTTAC 353
QY      114 GluAapGlyCySpheIleTyTrLyVallyAapheIleGlyValaAapheProSerAapGly 133

```

Db	354	GAGGGTTCCCATATCAAGCCGACATGAAGCTGATGGAGACGGTTTCCCTGACGACGGC	413
		:::	
Qy	134	ProValmetGlnLysThrMetGlyTTrpGlnIAser---	ThrgIArgLeuTyrPro 152
		:::	
Db	414	CCACTCATGACCCACCGACGATTTGTGACACGAGACGGCTGCGTGTCCAGAAAGACGTATCTT	473
Qy	153	ArgAepGlyValLeuLysGlyGlnIleHisValAlaLeuLysLeuLysAepGly-----	170
		:::	
Db	474	AACGACAAACCAATCTGCGACAGCTTGCATCGAGGTATTAACAATCTGCAGAAAGGAGACGC	533
Qy	171	-----GlyHisTyrIleuValGlu-----	PelySerIleTyrMet 182
		:::	
Db	534	TACAGGGCCCGAGTGTGAGCCACTACATCTTCCGACAAAGCCCTTTACAGCCGATCTATG	593
Qy	183	AlAllysPProValGlnLeuProGlyTyrTyrTyrValAepSerLysLeuAepIleThr	202
		:::	
Db	594	AAGAGACGACGGCTG-----TTCCGTGTCACGGAAGTGCACAGTAAAG	635
Qy	203	SerHisAnGluAepTyrThrIleValGlnGln	213
		:::	
Db	636	GCTTCCAAGACCGAAGTCACCTCTGACGAGAGG	668
RESULT 15			
LOCUS	AL004652	790 bp	mRNA
DEFINITION	DKRZp34P092_r1_434 (synonym: htes3) Homo sapiens CDNA clone		
VERSION	DKRZp34P092_5', mRNA sequence.		
KEYWORDS	AL004652.1 GI:5432867		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
REFERENCE	1 (bases 1 to 790)		
AUTHORS	Ansorge,W., Benez,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.		
TITLE	EST (Ansorge, Benez, et al.)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: MIPS		
	MIPS		
	IngoIsaendter Landstr.1, D-85764 Neuherberg, Germany		
	This is the 5' sequence of the clone insert		
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer		
	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;		
	Heidelberg/Germany) within the CDNA sequencing consortium of the		
	German Genome Project.		
	No nt sequence available.		
	This clone (DKRZp34P092) is available at the RZPD in Berlin.		
	Please contact the RZPD: Reactionszentrum, Heubnerweg 6, 14059		
	Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.		
FEATURES	Location/Qualifiers		
source	1..790		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="DKRZp34P092"		
	/issue_type="cDNA"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/clone_1lb="434 (synonym: htes3)"		
	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,07e-14	Length:	790
Score:	209.50	Matches:	55
Percent Similarity:	47.25%	Conservative:	48
Best Local Similarity:	25.23%	Mismatches:	102
Query Match:	17.26%	Indels:	13

```

DB: 1 Gaps: 6
US-10-006-922A-12 (1-225) x AL044652 (1-790)
QY 16 ValArgMetGluGlyThrValAlaMetGlyHisGluPheGluIleGluGlyGluGlyGly 35
Db 32 GTGAAATCGATGGCGATGTAATGGGACAAATTTTGTCTGACGGAGGAGGTAAGGT 91
QY 36 ArgProGlyGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuPro 55
Db 92 GATGCCACATGACGAAAGCTGACCTGTAAATTCATCTGCACCACT---GGAAAGCTCCCT 148
QY 56 PheAlaTrpAspIleLeuSerProGlnPheGlnGlySerLysValThrValValHis 75
Db 149 GTGCCATGGCCAAACGCTGTCACTACCTTCACTATGGGCTGCAGAGCTTTCCAGATAC 208
QY 76 ProAlaAspIlePro----AspTyrLysLysLeuSerPheProGluGlyPheLysTrp 93
Db 209 CCAGACCATATGAAGCAGCAGTACTTTTTCAAAGACGCGCATCCCGAGGGCTATGTCAG 268
QY 94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerLeu 113
Db 269 GAGGAAACCATCTTTTTCAAAGATACGGAACTACAAAGACCCGCGCTGAAGTCAAGTTC 328
QY 114 GlnAspGlyCysPheIleTyrValLysValLysPheIleGlyValAsnPheProSerAspGly 133
Db 329 GAAGGTGACACCCGTGTGATATGATATCGAGCTGAAGGGCATTTGACTTAAAGAGATGGA 388
QY 134 ProValMetGlnLysLysThrMetGlyTrpGluAlaSerThrGlnArgLeuTyrProArg 153
Db 389 AACATTTCTCGGCACAAAG---CTGAAATACACTATTAATCCCAACATGTGTACATCATG 445
QY 154 AspGlyValLeuLysGlyLysIleHisLysValAlaLeuLysLeuLys-----Asp 169
Db 446 GCCGCAAGACAAAMAATATGCGATCAAGCTCAACTTCMAATGACAGACAAACATTGAGAT 505
QY 170 GlnGlyHisTyrLeuValGluPhe-----LysSerIleTyrMetAlaLysProVal 187
Db 506 GGATTCGTCGACGCTGGCCGACCATTAATCAACAGAACATCCCAATCGGCGACGGCCCTGTG 555
QY 188 GlnLeuProGlyLysTyrTyrValAspSerLysLeuAspIleThr-----SerHis 204
Db 566 CTCCTCCCAACAACACATTAATCTGTSCAACCCAGTGCCTGCGCTGTAAGATCCCAAGAA 625
QY 205 AsnGluAspTyrThrIleValGluGlnTyrGlnArgThrGlnGlyArgHis 222
Db 626 AAGAGAGACACATGATGCTGCTGTGAGATTGTGACCGCTGCTGGATCACACAT 679

```

Search completed: January 12, 2006, 12:23:29  
Job time : 3750 secs

**THIS PAGE BLANK (USP 10)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 12, 2006, 09:45:12 ; Search time 179 Seconds  
(without alignments)  
2234.365 Million cell updates/sec

Title: US-10-006-922a-12  
1214

Perfect score: 1214  
Sequence: 1 MRSSKVIKEFMRFKRMG.....EDYIVEQYERGRHHLFL 225

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 segs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DB=xlp  
-Q/cg2\_1/USPTO.epool\_p/US10006922/rnatc.10012006.162409.18947/app.query.fasta\_1.391  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rnt -MINMATCH=0.1 -LOOCL=0  
-LOOEXT=0 -UNITs=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORMext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922.0CGN.1.1.290 @rnatc.10012006.162409.18947 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

Issued Patents NA: \*  
1: /cg2\_6/ptodata/1/ina/1 COMB.seq: \*  
2: /cg2\_6/ptodata/1/ina/5 COMB.seq: \*  
3: /cg2\_6/ptodata/1/ina/6 COMB.seq: \*  
4: /cg2\_6/ptodata/1/ina/6 COMB.seq: \*  
5: /cg2\_6/ptodata/1/ina/H COMB.seq: \*  
6: /cg2\_6/ptodata/1/ina/PCUS COMB.seq: \*  
7: /cg2\_6/ptodata/1/ina/PP COMB.seq: \*  
8: /cg2\_6/ptodata/1/ina/RE COMB.seq: \*  
9: /cg2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	678	3	US-09-459-956-6
2	1214	100.0	859	3	US-09-866-538-11
3	1214	100.0	859	3	US-09-866-538-11
4	1210	99.7	723	3	US-10-152-296-1
5	1210	99.7	6984	3	US-10-001-189-45
6	729.5	60.1	699	3	US-09-459-956-5
7	681	56.1	860	3	US-10-244-779-1
8	667	54.9	801	3	US-09-459-956-7
9	495	40.8	690	3	US-09-459-956-2

10	486.5	40.1	1079	3	US-09-609-161B-15	Sequence 15, Appl
11	486.5	40.1	1079	3	US-09-626-581D-64	Sequence 64, Appl
12	486.5	40.1	1079	3	US-09-415-765B-64	Sequence 64, Appl
13	486.5	40.1	1079	3	US-09-626-580C-64	Sequence 64, Appl
14	486.5	40.1	1085	3	US-09-277-716-15	Sequence 15, Appl
15	486.5	40.0	696	3	US-09-459-956-4	Sequence 45, Appl
16	485.5	40.0	720	3	US-09-839-650-1	Sequence 1, Appl
17	485.5	40.0	720	3	US-10-652-703A-1	Sequence 1, Appl
18	485.5	40.0	720	3	US-10-021-818A-3	Sequence 3, Appl
19	477.5	39.3	696	3	US-09-459-956-3	Sequence 3, Appl
20	474.5	39.1	1021	3	US-09-839-650-2	Sequence 2, Appl
21	474.5	39.1	1021	3	US-10-652-703A-2	Sequence 2, Appl
22	474.5	39.1	1104	3	US-09-277-716-30	Sequence 30, Appl
23	474.5	39.1	1104	3	US-09-609-161B-10	Sequence 30, Appl
24	474.5	39.1	1279	3	US-09-277-716-31	Sequence 31, Appl
25	474.5	39.1	1279	3	US-09-609-161B-31	Sequence 31, Appl
26	468.5	38.6	1482	3	US-09-977-897-1	Sequence 1, Appl
27	464.5	38.3	720	3	US-10-021-818A-1	Sequence 1, Appl
28	218.5	18.0	1559	3	US-09-049-475-6	Sequence 6, Appl
29	215.5	17.8	717	3	US-09-023-946B-3	Sequence 3, Appl
30	215.5	17.8	4196	3	US-09-453-315-1	Sequence 1, Appl
31	215.5	17.8	4196	3	US-09-204-117B-1	Sequence 1, Appl
32	215.5	17.8	6232	3	US-09-796-575-1	Sequence 1, Appl
33	215.5	17.8	7455	3	US-09-220-557-19	Sequence 19, Appl
34	215.5	17.8	7455	3	US-10-219-227-19	Sequence 19, Appl
35	215.5	17.8	7686	3	US-09-502-710-23	Sequence 23, Appl
36	215.5	17.8	7686	3	US-09-502-710-26	Sequence 26, Appl
37	215.5	17.8	7686	3	US-09-502-711-23	Sequence 23, Appl
38	215.5	17.8	7686	3	US-09-502-711-26	Sequence 26, Appl
39	215.5	17.8	7686	3	US-09-565-616A-2	Sequence 2, Appl
40	215.5	17.8	7687	3	US-09-502-710-24	Sequence 24, Appl
41	215.5	17.8	7687	3	US-09-502-711-24	Sequence 24, Appl
42	212.5	17.5	717	2	US-08-818-604-31	Sequence 31, Appl
43	212.5	17.5	717	3	US-09-346-946-31	Sequence 31, Appl
44	212.5	17.5	717	3	US-09-023-946B-21	Sequence 21, Appl
45	212.5	17.5	764	2	US-08-818-604-30	Sequence 30, Appl

#### ALIGNMENTS

RESULT 1  
US-09-459-956-6  
Sequence 6, Application US/0945956  
Patent No. 6342379  
GENERAL INFORMATION:  
APPLICANT: Teien, Roger Y.  
TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
TITLE OF INVENTION: OPTICAL METHODS  
FILE REFERENCE: REGEN1290-4  
CURRENT APPLICATION NUMBER: US/09/459, 956  
CURRENT FILING DATE: 1999-12-13  
PRIOR APPLICATION NUMBER: 08/765, 860  
PRIOR FILING DATE: 1999-05-08  
PRIOR APPLICATION NUMBER: 08/481, 977  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: PCT/US96/09652  
PRIOR FILING DATE: 1996-06-06  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 678  
TYPE: DNA  
ORGANISM: Discozoma sp  
US-09-459-956-6  
Alignment Scores:  
Pred. No.: 3.24e-160  
Score: 1214.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 3  
Length: 678  
Matches: 225  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0



```
SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-865-291-11

Alignment Scores:
Pred. No.: 4,72e-160 Length: 859
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-09-865-291-11 (1-859)
QY 1 MetATGSeSerLyAsnValIleLySgIuphMetArGpHeLyVaIaArgMetGluGly 20
DB 54 ATGAGGCTCTTCCAAAGAGTATCAAGAGGTTCAAGGTTTAAGGTTCCATGCAAGCA 113
QY 21 ThrValaEngIyHISgIuphGluIleGluGlyGluGlyArgProTyrgIuGly 40
DB 114 ACCGTCMAATGGGCAAGAGTTGAATAGAAAGGCAAGAGGAGGAGCCATACGAAGGC 173
QY 41 HIsaenThValLyLeuLyVaIThrLySgIyGlyProLeuProPheAlaTrpAapIle 60
DB 174 CACAATACCGTAAAGCTTAAGTAAAGCAAGGAGGAGCTTTGCCATTTCTGGAGATAT 233
QY 61 LeuSeProGlnPheGlnTyrgIySerLyVaIlyrValIyHIsaProIaAapIlePro 80
DB 234 TTGTACCAACAATTTCAGTAGAAGCAAGTATATGTAAGCAACCTTCCACATACCA 293
QY 81 AApTyLyLyLeuSeSerPheProGluGlyPheLySTrPGluArGyValMetAnPheGlu 100
DB 294 GACTATTAAGAGCTGCTCATTTCTCTGAAGATTTAAATGGAAAGGCTCATGAACCTTGA 353
QY 101 AApGlyLyIyValIyThrValIThrGlnAapSeSerLeuGlnAapGlyCySpHeIleTy 120
DB 354 GACGGGGGGTGGTCTTACTGTAACCCAGATTCAGATTCAGGATGGCTGTTTCATCTAC 413
QY 121 LyVaIyLyPheIleLyIyValaIaPheProSeArGpLyProValMetGlnLySlyThr 140
DB 414 AAGGTCMAATTCATTTGGCGTGAACCTTCTCCGATGACCTGTATAGCAAGAAAGACA 473
QY 141 MetGlyTrPGluAlaSeerThrgIuArGLeuTyrProArGpLyValIleuLySgIyGlu 160
DB 474 ATGGGGTGGGAAGCCGACATGAGCTTTGTATCCCTCGATGGCGGTGTTGAAGAGAG 533
QY 161 ILeHIsLyAlaLeuLyLeuLySaApGlyIyHIsTyrLeuValGluPheLySeriIe 180
DB 534 ATTCATTAAGGCTCTGAAGCTGAAGACGGGTGTCATTACTAGTTCAATTCAAAGATAT 593
QY 181 TyrMetAlaLyLySProValGlnLeuProGlyTyTyTyTyValaIaPseerLySleuAaP 200
DB 594 TACATGGCAAGAGGCTGTGAGCTACCAAGGCTACTATGTATGACTTCCAAAGTGAAT 653
QY 201 ILeThSeerHIsaEngIyAapTyrThrIleValIyGlnTyrgIuArGThrgIuGlyArg 220
DB 654 ATTAACAAGCAACAAGAGCTATACATCGTTGAGACAGTAAAGAACCGAGGAGCG 713
QY 221 HIsHIsleuPheLeu 225
DB 714 CACCATCTGTCTT 728

RESULT 4
US-10-152-296-1
; Sequence 1, Application US/10152296
; Patent No. 6723537
; GENERAL INFORMATION:
```

```
APPLICANT: Peelle, Beau
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells
; FILE REFERENCE: 021044-000110US
; CURRENT APPLICATION NUMBER: US/10/152,296
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/291,871
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mammalian
; OTHER INFORMATION: codon-optimized variant (DSRD) of Discosoma sp.
; OTHER INFORMATION: "red" red fluorescent protein (RFP)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(723)
; OTHER INFORMATION: DSRD
US-10-152-296-1

Alignment Scores:
Pred. No.: 1.31e-159 Length: 723
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-152-296-1 (1-723)
QY 1 MetATGSeSerLyAsnValIleLySgIuphMetArGpHeLyVaIaArgMetGluGly 20
DB 4 GTGGCTCTCTCCAAAGAGCTATCAAGAGGTTCAAGCGCTTCAAGGTGGCATGAGAGGC 63
QY 21 ThrValaEngIyHISgIuphGluIleGluGlyGluGlyArgProTyrgIuGly 40
DB 64 ACCGTGAAGGCAAGAGTTCAGATCGAGGCGAGGAGGAGGCGCCCTACAGAGGC 123
QY 41 HIsaenThValLyLeuLyVaIThrLySgIyGlyProLeuProPheAlaTrpAapIle 60
DB 124 CACAACACCGTAAAGCTGAAGGTGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 183
QY 61 LeuSeProGlnPheGlnTyrgIySerLyVaIlyrValIyHIsaProIaAapIlePro 80
DB 184 CTGTCCCCCAATTCAGTACGGCTCAAGGTGTAAGTAAACACCCCGACATCCCC 243
QY 81 AApTyLyLyLeuSeSerPheProGluGlyPheLySTrPGluArGyValMetAnPheGlu 100
DB 244 GACTCAAGAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAACCTTCAG 303
QY 101 AApGlyLyIyValIyThrValIThrGlnAapSeSerLeuGlnAapGlyCySpHeIleTy 120
DB 304 GACGGGGGGTGGTGGCTGACCCGTAACCAAGCTCTTCCGTAAGAGCGCTGCTCATCTAC 363
QY 121 LyVaIyLyPheIleGlyValaIaPheProSeArGpLyProValMetGlnLySlyThr 140
DB 364 AAGGTAAAGTTCATCGGCTGAACCTTCCCTCGAAGGCGCCGCTGAAGAGAGACC 423
QY 141 MetGlyTrPGluAlaSeerThrgIuArGLeuTyrProArGpLyValIleuLySgIyGlu 160
DB 424 ATGGGGTGGAGGCTTCCACGAGCGCTGTACCCCGAGAGCGGTGTAAGGCGAG 483
QY 161 ILeHIsLyAlaLeuLyLeuLySaApGlyIyHIsTyrLeuValGluPheLySeriIe 180
DB 484 ATTCACAAAGGCTCTGAAGCTGAAGAGCGGAGGAGCTACTAGTGAAGTTCAAAGATAC 543
QY 181 TyrMetAlaLyLySProValGlnLeuProGlyTyTyTyTyValaIaPseerLySleuAaP 200
DB 544 TACATGGCAAGAGCGCTGTGAGCTGCCGGCTACTATTAAGTGAATCCAAAGCTGAGAC 603
```

Qy 201 lIeThSerHisAenGluAapTyThrIleValGluGlnTyrgIuAargThrGluGlyArg 220  
|  
Db 604 ATCACTCCCAAGAGAGACTACACATCGTAGAGCAAGCGCACCGAGGGCCGC 663  
Qy 221 HisHisLeuPheLeu 225  
|  
Db 664 CACCACCTGTTCTCG 678  
RESULT 5  
US-10-001-189-45  
; Sequence 45, Application US/10001189  
; Patent No. 6962810  
; GENERAL INFORMATION:  
; APPLICANT: FRASER JR., MALCOLM J.  
; APPLICANT: LI, XU  
; APPLICANT: BEAM, TERESA  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING  
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION  
; FILE REFERENCE: 835910-92098  
; CURRENT APPLICATION NUMBER: US/10/001,189  
; CURRENT FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/244,984  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/244,677  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 45  
; LENGTH: 6984  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE: Artificial Sequence  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: pbXp3-Derived-orf sequence  
US-10-001-189-45  
Alignment Scores:  
Pred. No.: 4.87e-158 Length: 6984  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 3 Gaps: 0  
US-10-006-922a-12 (1-225) x US-10-001-189-45 (1-6984)  
Qy 1 MeArGSeSerLySaeNValIleLyGluPheMeArGpHeLyVaIArgMeGluGly 20  
|  
Db 2974 GTGGCTCTCCAGAGAGCTCATCAAGAGTTTCATGCCCTTCAAGGTGCGCATGAGGGC 3033  
Qy 21 ThrValaAngIyHlBgIuPheGluIleGluGlyGluGlyArgProTyrgIuGly 40  
|  
Db 3034 ACCGGAAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 3093  
Qy 41 HisAenThrValLySLeuLyVaIThrLyGlyGlyProLeuProPheAlaTpaAspIle 60  
|  
Db 3094 CACAACACCGGAGAGCTGAAGGTGACCAAGGGCGGCCCTTGGCTTGGAGACATC 3153  
Qy 61 LeuSerProGlnPheGlnTyrgIySerLySaeNValTyrgValLySHlAProAlaAspIlePro 80  
|  
Db 3154 CTGGTCCCCCGATTCCATGACGGCTCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 3213  
Qy 81 AspTyrgLySLeuSerPheProGluGlyPheLySTpGluArgValMetAenPheGlu 100  
|  
Db 3214 GACTACAAAGAGGCTGCTCTCCCGAGGGCTTCAAGTGGGAGCGCGTGTATGAATTCAG 3273  
Qy 101 AspGlyGlyValIleThrValThrGlnAenSerSerLeuGlnAenGlyCyAProPheIleTy 120  
|  
Db 3274 GACGGCGCGGTGTACCGGTGACCGAGATCTCTCCGTGCAAGAGGGCTGCTTCAATTCAC 3333  
Qy 121 LySValLySPhelIeGlyValAenPheProSerAenGlyProValMetGlnLySLeuThr 140

Db 3334 AAGGTAAAGTTTCATCGGCGTGAATCTCCCTCCGACGGCCCCGTAATGCAAGAAAGACC 3393  
Qy 141 MeGlyTyrgIuAaSerThrgIuAargLeuTyrgProArgAapGlyValIleuLyGlyGlu 160  
|  
Db 3394 ATGGGCTGGAGGGCTTCCACCGAGCGCTGTACCCCGCACCGCGTGTGAAGGGCGAG 3453  
Qy 161 lIeHisLySAlaLeuLySLeuLyAapGlyGlyHisTyrgLeuValGluPheLySerIle 180  
|  
Db 3454 ATCCCAAGGCCCTGAAGCTGAAGAGACGGCGGCACCTACCTGGTGTGAATTCATTCATC 3513  
Qy 181 TyrMeAlaLySlySProValGlnLeuProGlyTyrgTyrgValAaPSeLySLeuAap 200  
|  
Db 3514 TACATGGCCCAAGAAAGCCCGTGCAGCTGCCGGCTACTACTACGTGAGACTCCAACTGAC 3573  
Qy 201 lIeThSerHisAenGluAapTyThrIleValGluGlnTyrgIuAargThrGluGlyArg 220  
|  
Db 3574 ATCACTCCCAAGAGAGACTACACATCGTAGAGCAATGAGCGCACCGAGGGCCGC 3633  
Qy 221 HisHisLeuPheLeu 225  
|  
Db 3634 CACCACCTGTTCTCG 3648  
RESULT 6  
US-09-459-956-5  
; Sequence 5, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger Y.  
; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; TITLE OF INVENTION: OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: FCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 699  
; TYPE: DNA  
; ORGANISM: *Drosophila striata*  
US-09-459-956-5  
Alignment Scores:  
Pred. No.: 2.87e-92 Length: 699  
Score: 729.50 Matches: 133  
Percent Similarity: 76.00% Conservative: 38  
Best Local Similarity: 59.11% Mismatches: 53  
Query Match: 60.09% Indels: 1  
DB: 3 Gaps: 1  
US-10-006-922a-12 (1-225) x US-09-459-956-5 (1-699)  
Qy 1 MeArGSeSerLySaeNValIleLyGluPheMeArGpHeLyVaIArgMeGluGly 20  
|  
Db 1 ATGAGTTGTTCCAAAGAGGTGATCAAGAGAAATGTGATCGATCTTCATCTGGAAGCA 60  
Qy 21 ThrValaAngIyHlBgIuPheGluIleGluGlyGluGlyArgProTyrgIuGly 40  
|  
Db 61 ACGTTCAATGGGCACTCTTGAATATAAGGCAAGGAAAGGACAGCTTAATGAAGGC 120  
Qy 41 HisAenThrValLySLeuLyVaIThrLyGlyGlyProLeuProPheAlaTpaAspIle 60  
|  
Db 121 ACCAATACGTTACGCTCGAGGTTCACAAAGGTGAGACTCTGCGCAATTTGGTCATAT 180  
Qy 61 LeuSerProGlnPheGlnTyrgIySerLySaeNValTyrgValLySHlAProAlaAspIlePro 80

Db 181 TTGTCACCAATTTTCAGATGGAACAAGGCAATTTGTCACACCTGACACATACAT 240  
Qy AaprrlylvleuSerPheProgluGlyPheylstPgluArgValMetanPheglu 100  
Db 241 GATTATCTAAAGCTTCATTTCCGAGGATATACATGGACGGCTTCATGACATTTGA 300  
Qy 101 AaprglyglValValThrValThrGlnAspSerSerLeuGlnAspGlyCyapheleTy 120  
Db 301 GACGGTGGCTTGCTGTGATATCCAAATGATATACATTTGACAGGCAACTGTTTCTACTAC 360  
Qy 121 LysValIlyPheIleGlyValAsnPheProSerAspGlyProValMetGlnIlystThr 140  
Db 361 GACATCAAGTTCACGGCTTGAACTTTCTCCAAATGACCCGTTGTGCAAGAAAGACA 420  
Qy 141 MetGlyTPGluAlaSerThrGluArgLeuTyProArgAspGlyValIleuysGlylu 160  
Db 421 ACTGGCTGGAAACCGAGCACTAGAGGTTGTATCTCGTATGCTGTGTGATAGAGAC 480  
Qy 161 IleIstlyAlaIleuIlyleuLysAspGlyGlyIstYrLeuValGluPheIlySerIle 180  
Db 481 ATCCATCATGCTCTGACATGTAAGAGGTGGTCAATTAAGCATGACATTAACCTGTT 540  
Qy 181 TyMetAlaLysLys---ProValGlnLeuProGlyTyTyTyTyValAspSerIlyleu 199  
Db 541 TACAGGGCCAAAGAGCCGCTTGAAAGATGACAGGGTATCACTATGTGTACACCAACTG 600  
Qy 200 AaprlIethSerHisAsnGlnAspTyThrIleValGluGlnTyArgThrGluGly 219  
Db 601 GTTATATGAGAACACCAAGAAATTCATGAAGTTGAGAGCATGAATCCGCTTGCA 660  
Qy 220 ARGHleIleuphe 224  
Db 661 CCCACCATCCGTTT 675

RESULT 7  
US-10-244-779-1  
Sequence 1, Application US/10244779  
Patent No. 6933375

GENERAL INFORMATION:  
APPLICANT: Falkowski, Paul  
APPLICANT: Sun, Yi  
APPLICANT: Gorbunov, Maxim  
APPLICANT: Wyman, Kevin  
APPLICANT: Chen, Yi-Bu  
TITLE OF INVENTION: mcrp Encoding Nucleic Acids,  
TITLE OF INVENTION: Polypeptides, Antibodies and Methods of Use Thereof  
FILE REFERENCE: Rut 00-0023US  
CURRENT APPLICATION NUMBER: US/10/244,779  
CURRENT FILING DATE: 2002-12-19  
PRIOR FILING DATE: 2001-09-14  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 860  
TYPE: DNA  
ORGANISM: Montastrea cavernosa  
US-10-244-779-1

## Alignment Scores:

Pred. No.: 2,52e-85 Length: 860  
Score: 681.00 Matches: 123  
Percent Similarity: 73.95% Conservative: 36  
Best Local Similarity: 57.21% Mismatches: 56  
Query Match: 56.10% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-244-779-1 (1-860)

Qy 6 AsnValIleuysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGlyHis 25  
Db 138 AGGTGATTAATCATGATGAAATATCAAGCTGCGTATGACGGCATTTGTAACGGGCAC 197

Qy 26 GluphegluIleGluGlyGluGlyArgProTyrgIuGluIleAsnThrValLys 45  
Db 198 AAGTTCATGATTAACGAGGAGGTAAGCAAGCTTTGAGGGAACACATATATA 257  
Qy 46 LeuysValIthrGlyGlyProleuProPheAlaTPaPilleuSerProGlnPhe 65  
Db 258 CTTAAAGTCAAGAAAGCGGACCTCTGCTTTGCTTTCGATTCGATTCACACATTT 317  
Qy 66 GlnTyrgIyseryValTyValIlyshiProAlaAspIleProAspTyTyLysLys 85  
Db 318 CAGTACGGCAACAGGGTATTCACCAATATACCAAAAGACATAACGACATTTTCAAGANG 377  
Qy 86 SerPheProGluGlyPheLysThrGluArgValMetanPheGluAspGlyValVal 105  
Db 378 TCGTTTCTGAGGGGTATTCCTGGAAAGACATGACTTGTGAAGACAGGCGCTTGC 437  
Qy 106 ThrValIthrGlnAspSerSerLeuGlnAspGlyCyapheleTyTyValIlyAspHele 125  
Db 438 ACCGTCAAGGACATATAAGTTGGAAGCGACCTTTTCTACGAATTCATTTAT 497  
Qy 126 GlyValAsnPheProSerAspGlyProValMetGlnIlystThrMetGlyTPGluAla 145  
Db 498 GGTGGAATTTCCCTCCAGTGTCCAGTTATGCAGAAAGACCTGAAATGAGGACCA 557  
Qy 146 SerThrGluArgLeuTyProArgAspGlyValIleuysGlyGluIleIstlyAlaLys 165  
Db 558 TCCACTGGAATATATAGCTGTGATGATGATGCTGAGGAGTTATGACAGACCTG 617  
Qy 166 LysLeuLysAspGlyGlyIstYrLeuValGluPheIlySerIleTyMetAlaLysLys 185  
Db 618 TTGCTTGAAAGGGAATTAACATCACCGATCTACAGAAATCTTACGGGCAAGAG 677  
Qy 186 ProValGlnLeuProGlyTyTyTyTyValAspSerIlyleuAspIleThrSerHis 205  
Db 678 GGTGCTGTGCTCCAGAAATATCATCTTGTGACCAACGAAATGAAATCTGAGCATGAC 737

US-09-459-956-7

Sequence 7, Application US/0945956  
Patent No. 6342379  
GENERAL INFORMATION:  
APPLICANT: Teisen, Roger Y.  
TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
TITLE OF INVENTION: OPTICAL METHODS  
FILE REFERENCE: REGEN1290-4  
CURRENT APPLICATION NUMBER: US/09/459,956  
CURRENT FILING DATE: 1999-12-13  
PRIOR APPLICATION NUMBER: 08/765,860  
PRIOR FILING DATE: 1999-05-08  
PRIOR APPLICATION NUMBER: 08/481,977  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: PCT/US96/09652  
PRIOR FILING DATE: 1996-06-06  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 801  
TYPE: DNA  
ORGANISM: Clavularia sp  
US-09-459-956-7

## Alignment Scores:

Pred. No.: 2,06e-83 Length: 801  
Score: 667.00 Matches: 121  
Percent Similarity: 73.76% Conservative: 42  
Best Local Similarity: 54.75% Mismatches: 58  
Query Match: 54.94% Indels: 0  
DB: 3 Gaps: 0



/ PRIOR FILING DATE: 1999-03-26  
/ PRIOR APPLICATION NUMBER: 60/102,939  
/ PRIOR FILING DATE: 1998-10-01  
/ PRIOR APPLICATION NUMBER: 60/089,367  
/ PRIOR FILING DATE: 1998-06-15  
/ PRIOR APPLICATION NUMBER: 60/079,624  
/ PRIOR FILING DATE: 1998-03-27  
/ NUMBER OF SEQ ID NOS: 32  
/ SOFTWARE: Patentin Ver. 2.0  
/ SEQ ID NO 15  
/ LENGTH: 1079  
/ TYPE: DNA  
/ ORGANISM: Renilla mulleri  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (259)..(975)  
/ OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)  
US-09-609-161B-15

Alignment Scores:  
Pred. No.: 6,71e-58 Length: 1079  
Score: 486.50 Matches: 97  
Percent Similarity: 64.09% Conservative: 44  
Best Local Similarity: 44.09% Mismatches: 72  
Query Match: 40.07% Indels: 7  
Gaps: 3

US-10-006-922a-12 (1-225) x US-09-609-161B-15 (1-1079)

QY 2 ArgSerSerlybAnValIlelyb-----GluphemeTarGhPheLyVal 16  
DB 256 AAGATGAGTAAACAAATATGAAAGAACACTGTCTTACAAAGAGTAATGCGTAAAGTA 315  
QY 17 ArgMeGluGlyThrValaAngLyHISgluPheGluileGluGlyGluGlyArg 36  
DB 316 AATCTGGAAGAAATTGTAAACAACCATGTTTACATGAGAGGCTCCGCAAGGAAT 375  
QY 37 ProTyGluGlyHISanThrValLyLeuLyValThrLyseGlyProLeuProPhe 56  
DB 376 ATTTATTCGGCAATCAACTGCTTCAAGTTCGATGCAAGAGGGCCCACTGCCCTTT 435  
QY 57 AlaTrpAspIleLeuSerProGlnPheGlnTyGlySerLybValTyValLybHISPro 76  
DB 436 GCATTGATATGTCACAGAGCTTTCAATATGCGCAACGCTTCTTCAAGAAATATCCG 495  
QY 77 AlaAspIleProAspTyLybLybLeuSerPheProGluGlyPheLybTPGluArgVal 96  
DB 496 AATGATATATCAAGATTTATTTATATCAATCATTTCCACAGAGATTTATGATGAACGACA 555  
QY 97 MetAsnPheGluAspGlyGlyValaThrValThrGlnAspSerSerLeuGlnAspGly 116  
DB 556 TTACGTTACGAAGATGCGGAGCTTGTGAATTCGTCAGATTAATTAATATGAAGAC 615  
QY 117 CybPheIleTyLybValLybPheIleGlyValaAsnPheProSerAspGlyProValMet 136  
DB 616 AAGTTCGCTTACAGAGAGATACAAAGGTATGTAATCTCCAGATGATGTCGCCGTATG 675  
QY 137 GlnLybSerThrMetGlyTPGluAlaSerThrGluArgLeuTyProAspAspGlyVal 156  
DB 676 CAGAAAGCTATCTTAGAATAGAGCTTCAATTGAAAGCCATGATGAATATAGGCGTC 735  
QY 157 LeuLyseGlyLulleHISLybAlaLeuLybLeuLybAspGlyGlyHISLyLeuValGlu 176  
DB 736 TTGGTGGCGGAAGTAATCTTGTCTATTAATTAACCTGGGAATATTAATTCATGTCAC 795  
QY 177 PheLybSerIleTyLybMetAlaLybLybProVal---GlnLeuProGlyTyTyTyVal 195  
DB 796 ATGAAACATTAATGAAGTCGAAAGGTGTATGAAGAGTTCCTTCGATCATTTAAT 855  
QY 196 ArgSerLybLeuAspIleThrSerHISanGluAspTyThrIleValGluGlnTyGlu 215  
DB 856 CACATCGTTTGGAAAGACT--TACGTAGAAAGCGGGGGTTCGTTGAACAGCATGAG 912

## RESULT 11

US-09-626-581D-64  
/ Sequence 64, Application US/09626581D  
/ Patent No. 6548249  
/ GENERAL INFORMATION:  
/ APPLICANT: Anderson, David  
/ TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide  
/ FILE REFERENCE: A-66900-3/RMS  
/ CURRENT APPLICATION NUMBER: US/09/626,581D  
/ PRIOR FILING DATE: 2000-07-27  
/ PRIOR APPLICATION NUMBER: 09/169,015  
/ PRIOR FILING DATE: 1998-10-08  
/ PRIOR APPLICATION NUMBER: 09/415,765  
/ NUMBER OF SEQ ID NOS: 65  
/ SOFTWARE: Patentin Ver. 2.0  
/ SEQ ID NO 64  
/ LENGTH: 1079  
/ TYPE: DNA  
/ ORGANISM: Renilla mulleri  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (259)..(975)  
/ OTHER INFORMATION:  
US-09-626-581D-64

Alignment Scores:  
Pred. No.: 6,71e-58 Length: 1079  
Score: 486.50 Matches: 97  
Percent Similarity: 64.09% Conservative: 44  
Best Local Similarity: 44.09% Mismatches: 72  
Query Match: 40.07% Indels: 7  
Gaps: 3

US-10-006-922a-12 (1-225) x US-09-626-581D-64 (1-1079)

QY 2 ArgSerSerlybAnValIlelyb-----GluphemeTarGhPheLyVal 16  
DB 256 AAGATGAGTAAACAAATATGAAAGAACACTGTCTTACAAAGAGTAATGCGTAAAGTA 315  
QY 17 ArgMeGluGlyThrValaAngLyHISgluPheGluileGluGlyGluGlyArg 36  
DB 316 AATCTGGAAGAAATTGTAAACAACCATGTTTACATGAGAGGCTCCGCAAGGAAT 375  
QY 37 ProTyGluGlyHISanThrValLyLeuLyValThrLyseGlyProLeuProPhe 56  
DB 376 ATTTATTCGGCAATCAACTGCTTCAAGTTCGATGCAAGAGGGCCCACTGCCCTTT 435  
QY 57 AlaTrpAspIleLeuSerProGlnPheGlnTyGlySerLybValTyValLybHISPro 76  
DB 436 GCATTGATATGTCACAGAGCTTTCAATATGCGCAACGCTTCTTCAAGAAATATCCG 495  
QY 77 AlaAspIleProAspTyLybLybLeuSerPheProGluGlyPheLybTPGluArgVal 96  
DB 496 AATGATATATCAAGATTTATTTATATCAATCATTTCCACAGAGATTTATGATGAACGACA 555  
QY 97 MetAsnPheGluAspGlyGlyValaThrValThrGlnAspSerSerLeuGlnAspGly 116  
DB 556 TTACGTTACGAAGATGCGGAGCTTGTGAATTCGTCAGATTAATTAATATGAAGAC 615  
QY 117 CybPheIleTyLybValLybPheIleGlyValaAsnPheProSerAspGlyProValMet 136  
DB 616 AAGTTCGCTTACAGAGAGATACAAAGGTATGTAATCTCCAGATGATGTCGCCGTATG 675  
QY 137 GlnLybSerThrMetGlyTPGluAlaSerThrGluArgLeuTyProAspAspGlyVal 156  
DB 676 CAGAAAGCTATCTTAGAATAGAGCTTCAATTGAAAGCCATGATGAATATAGGCGTC 735  
QY 157 LeuLyseGlyLulleHISLybAlaLeuLybLeuLybAspGlyGlyHISLyLeuValGlu 176  
DB 736 TTGGTGGCGGAAGTAATCTTGTCTATTAATTAACCTGGGAATATTAATTCATGTCAC 795

Qy	177	PheIySeSerIleYrMetAlaIySvProVal---	GlnLeuProGlyYrYrYrYrVal	195
Db	796	ATGAAACATTATATGAATCGGAAGGTGATGTAAGGAGTTTCTTCGTATCACTTATATT		855
Qy	196	ArgSerIleuAspIleThSerHisAngIuAspYrTrHrIleValGluGlnTrGlu		215
Db	856	CAACATCGTTTGGAAGAGACT--TACGTAGAAGACGGGGGGGTTCTGTGAACACATGAG		912
RESULT 12				
US-09-415-765B-64				
; Sequence 64, Application US/09415765B				
; Patent No. 6548632				
; GENERAL INFORMATION:				
; APPLICANT: Anderson, David				
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide				
; FILE REFERENCE: Libraries				
; CURRENT APPLICATION NUMBER: US/09/415,765B				
; PRIOR APPLICATION NUMBER: 1999-10-08				
; PRIOR FILING DATE: 1998-10-08				
; NUMBER OF SEQ ID NOS: 65				
; SOFTWARE: PatentIn Ver. 2.0				
; SEQ ID NO 64				
; LENGTH: 1079				
; TYPE: DNA				
; ORGANISM: Renilla muelleri				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: (259)..(975)				
; OTHER INFORMATION:				
US-09-415-765B-64				
Alignment Scores:				
Pred. No.: 6,71e-58 Length: 1079				
Score: 486.50 Matches: 97				
Percent Similarity: 64.09% Conservative: 44				
Best Local Similarity: 44.09% Mismatches: 72				
Query Match: 40.07% Indels: 7				
DB: 3 Gaps: 3				
US-10-006-922A-12 (1-225) x US-09-415-765B-64 (1-1079)				
Qy	2	ArgSerSerIleYrAsnValIleIyS-----GluPheMetArgPheIySVal	16	
Db	256	AAGATGAGTAACCAATATTTGAAGAACACTTGTTCACAAAGAGTAATGTCGTATTAAGTA	315	
Qy	17	ArgMetGluGlyThrValaAngIyHisGluPheGluIleGluGlyValGluGlyArg	36	
Db	316	AATCTGGAAGAAATTGTAAACAACATGTTTTACAAATGAGGGGTTGGCGCAAGGGAAT	375	
Qy	37	ProIyrgLugIyHisAsnThrValIySLeuIySValThrIySgIyIyProIeuProPhe	56	
Db	376	ATTTTATTTGGCAATCACTACGTTGAGATGTCGTCAAGAAAGGGCCCACTGCTTTT	435	
Qy	57	AlaTrpAspIleIeuSerProGlnPheGlnTrYrGlySerIySValIyYrValIyHisIlePro	76	
Db	436	GCAATTGATATTGTGTCCACACCTTTTCAMTATGGCAACCGTACTCTTACACGAATATCCG	455	
Qy	77	AlaAspIleProAspTrIyIySLeuSerPheProGluIyPheIySTrpGluArgVal	96	
Db	496	AATGATATATTCAGATTATTTTATACATCATTTCCACGACGAGATTATGATGAACGAACA	555	
Qy	97	MetAsnPheGluAspGlyIyValValThrValThrGlnAspSerSerIeuGlnAspGly	116	
Db	556	TTTACGTTCAGGAAGATGCGGCACTGTGTGAATTCGTCAGATATTAATTTATAGAAGAC	615	
Qy	117	CysPheIleTrIyIySValIySPhelIeGlyValaAsnIleProSerAspGlyProValMet	136	
Db	616	AAGTTGTCTACAGATGGAATCAACAAGGTACTTACTTCCACAGATGATGTGCTCCGTATG	675	
Qy	137	GlnIyIySLeuThrMetGlyTrpGluIaSerThrGluArgLeuYrProArgAspGlyVal	156	

```

Db      676 CAGAAAGACTATCTTAGAATATAGAGCCCTTCATTGAAAGCAGTACATGAAATATAGCGCC 725
Qy      157 Leuysgylguillehlsylnaleuysleuysaapglglyhlsfryleuvalglu 176
Db      736 TTGGTCGCGAAGTAATCTTGTCATATAACTAACTGCGAATAATATATATCATGTCAC 795
Qy      177 Phelyseriletyrmetallalyseproval---Glnleuproglytyrtyrval 195
Db      796 ATGAAAACATTAAAGAACTCGAAGAGGTAGTAAAGAACTTCCCTTCATCATTTTATT 855
Qy      196 AepserleleAepillethserisangniAapTyrrthrlvalglugintyrglu 215
Db      856 CAACATCTTTGGAAAAGACT---TACGTAGAAAGACGGGGGTTCTTGAACGATGAG 912

RESULT 13
US-09-626-580C-64
; Sequence 64, Application US/09626580C
; Patent No. 6562617
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: FUSIONS OF SCAFFOLD PROTEINS WITH RANDOM PEPTIDE LIBRARIES
; FILE REFERENCE: A-66900-2/RMS/AMS
; CURRENT APPLICATION NUMBER: US/09/626,580C
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Renilla muelleri
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (259)..(975)
; OTHER INFORMATION:
US-09-626-580C-64

Alignment Scores:
Pred. No.: 6.71e-58 Length: 1079
Score: 486.50 Matches: 97
Percent Similarity: 64.09% Conservative: 44
Best Local Similarity: 44.09% Mismatches: 72
Query Match: 40.07% Indels: 7
Db: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-626-580C-64 (1-1079)
Qy      2 ArgserSerlysaenVallelyse-----GluPhemeArGpHeuVal 16
Db      256 AAGATGAGTAAACAATATGTGAAGAACCTGTTTACAAGAAAGTATGTCGTATAAAGTA 315
Qy      17 ArgMetGluGlyThrValAsnGlyHlsGluPheGluIleGluGlyGluGlyGlyGly 36
Db      316 AATCTGGAAGAAATGTGTAACAACACATGTTTTTACAATAGSAGSGTTGCCGCAAGGAAT 375
Qy      37 ProTyrgluGlyHlsAsnThrVallyseuysValThrlyeGlyGlyProleuProphe 56
Db      376 ATTTTATTCGGCAATCAACGTGGTTCAGATTCGTGTACCAAGAGGGGCCCACTGCCCTTT 435
Qy      57 AlaTrpAplleleuSerProGlnPheGlnTyrglySerlyeValTyrrVallyshIAsPro 76
Db      436 GCATTGTGATATGTCGTACCAAGCTTTTCAATATGCAACCGTACTTTCGAAATATCCG 495
Qy      77 AlaAplleProAapTyrrlyeLyseuSerPheProGluGlyPhelysTrpGluatGval 96
Db      496 AATGATATATCAGATTATTTTATCAATCACTTTCCAGCAGGAGTTTATGTATGAAACAA 555
Qy      97 MetAsnPhelGluAapGlyValValThrValThrGlnApsSerSerleuGlnAapGly 116

```



```
DB 556 TTACGTTACGAAGATGCGCGACTTGTGAAATTCCTTCAGATATTAATTAATTAAGAC 615
QY 117 CyspHeileTyLyVaIllysheIlleglYalaAnpheProSeAaspGlyProValMet 136
DB 616 AAGTCGCTACAGATGGAATACAAAGGAGTAACTTCCAGATGATGTCCTCCATG 675
QY 137 GlnlyblystHmeGlyTTrpGlnlaseThrglUaArgLeuTyProArgaspGlyVal 156
DB 676 CAGAAAGACTATCTTGGAATAGAGCCTTATTGAAGCATGATGATGAATATGAGCGTC 735
QY 157 LeuLySGlyGluIlleHlyblyAlaLeuLyblybAaspGlyGlyHlystYrLeuValGlu 176
DB 736 TTGGTCGGCGAAGTAACTTGTCTATTAATACTTGGAATATTATTCATGTCAC 795
QY 177 PheLyseTleTyMeAlaLyblybProVal---GlnLeuProGlyTyTyTyVal 195
DB 796 ATGAAACATTATGATGAGTGAAGGTGATGAAGAGTTCCTTCGATCATTTTAT 855
QY 196 AapSerLyLeuAapIleThSerHlysaNgIuAapTyThrlEValGluGlnTyrglu 215
DB 856 CAACATCGTTTGAAAGAAGCT---TACGTAGAAGACGGGGGGTTCGTGAACGATGAG 912

RESULT 14
US-09-277-716-15
/ Sequence 15, Application US/09277716A
/ Patent No. 6232107
/ GENERAL INFORMATION:
/ APPLICANT: Bryan, Bruce
/ APPLICANT: Szent-Gyorgyi, Christopher
/ APPLICANT: PROLINE, LTD.
/ TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
/ CURRENT FILING DATE: 1999-03-26
/ EARLIER APPLICATION NUMBER: 60/102,939
/ EARLIER FILING DATE: 1998-10-01
/ EARLIER APPLICATION NUMBER: 60/089,367
/ EARLIER FILING DATE: 1998-06-15
/ EARLIER APPLICATION NUMBER: 60/079,624
/ EARLIER FILING DATE: 1998-03-27
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 15
/ LENGTH: 1079
/ TYPE: DNA
/ ORGANISM: Renilla mulleri
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (259)..(975)
/ FEATURE:
/ OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-277-716-15

Alignment Scores:
Pred. No.: 6,77e-58 Length: 1085
Score: 486.50 Matches: 97
Percent Similarity: 64.09% Conservative: 44
Best Local Similarity: 44.09% Mismatches: 72
Query Match: 40.07% Indels: 7
DB: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-277-716-15 (1-1085)
QY 2 ArgSerSerLybAanValIlleLyS---GlnPheMeTArgPheLybVal 16
DB 256 AAGATGATGAACAAGATATGAAAGACACTGTTTACAGAGAGTAAATGCTGATTAAGTA 315
QY 17 ArgMeGlyGluTyThrValaAnGlyHlyblybPheGluIlleglUgLyGluGlyArg 36
DB 316 AATCTGAAAGAAATGTGAACAACCAATGTTTACATGAGAGGGTTCGGCAAGGAAT 375
QY 37 ProTyGluGlyHlybAanThrValLybLeuLybValThrlEValGlyGlyProLeuProPhe 56
```

```
DB 376 ATTATTCGGCAATCACTGTTCAAGTTGCTGTCAGAAAGGGGCCCACTGCTTTT 435
QY 57 AlaTpaSpIleLeuSeProGlnPheGlnTyrglyblybValTyVallyHlybPro 76
DB 436 GCATTGATATGTTGTCAACAGCTTTTCATATATGCAACCGACTTTCACGAAATATCCG 495
QY 77 AlaAapIleProAapTyTyLyblybLeuLybPheProGluGlyPheLybTrpGluVal 96
DB 496 AATGATATATCATGATTTATTTATACATCATTTCCAGCAGAGATTTATGTATGACGAACA 555
QY 97 MetAanPheGluAaspGlyGlyValValThrValThrglnAapSerSerLeuGlnAaspGly 116
DB 556 TTACGTTACGAAGATGCGCGACTTGTGAATTCCTTCAGATATTAATTAATAGAACAC 615
QY 117 CyspHeileTyLyVaIllysheIlleglYalaAnpheProSeAaspGlyProValMet 136
DB 616 AAGTCGCTACAGATGGAATACAAAGGAGTAACTTCCAGATGATGTCCTCCATG 675
QY 137 GlnlyblystHmeGlyTTrpGlnlaseThrglUaArgLeuTyProArgaspGlyVal 156
DB 736 TTGGTCGGCGAAGTAACTTGTCTATTAATACTTGGAATATTATTCATGTCAC 795
QY 157 LeuLySGlyGluIlleHlyblyAlaLeuLyblybAaspGlyGlyHlystYrLeuValGlu 176
DB 796 TTGGTCGGCGAAGTAACTTGTCTATTAATACTTGGAATATTATTCATGTCAC 795
QY 177 PheLyseTleTyMeAlaLyblybProVal---GlnLeuProGlyTyTyTyVal 195
DB 796 ATGAAACATTATGATGAGTGAAGGTGATGAAGAGTTCCTTCGATCATTTTAT 855
QY 196 AapSerLyLeuAapIleThSerHlysaNgIuAapTyThrlEValGluGlnTyrglu 215
DB 856 CAACATCGTTTGAAAGAAGCT---TACGTAGAAGACGGGGGGTTCGTGAACGATGAG 912

RESULT 15
US-09-459-956-4
/ Sequence 4, Application US/09459956
/ Patent No. 6342379
/ GENERAL INFORMATION:
/ APPLICANT: Teien, Roger Y.
/ APPLICANT: Gonzalez, Ili, Jesus E.
/ TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
/ TITLE OF INVENTION: OPTICAL METHODS
/ FILE REFERENCE: REGEN1290-4
/ CURRENT APPLICATION NUMBER: US/09/459,956
/ CURRENT FILING DATE: 1999-12-13
/ PRIOR APPLICATION NUMBER: 08/765,860
/ PRIOR FILING DATE: 1999-05-08
/ PRIOR APPLICATION NUMBER: 08/481,977
/ PRIOR FILING DATE: 1995-06-07
/ PRIOR APPLICATION NUMBER: PCT/US96/09652
/ PRIOR FILING DATE: 1996-06-06
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ TYPE: DNA
/ ORGANISM: Zoanthus sp
US-09-459-956-4

Alignment Scores:
Pred. No.: 4.6e-58 Length: 696
Score: 485.50 Matches: 95
Percent Similarity: 64.71% Conservative: 37
Best Local Similarity: 46.57% Mismatches: 67
Query Match: 39.99% Indels: 5
DB: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-459-956-4 (1-696)
QY 1 MetArgSerSerLybAanValIlleLySGluPheMeTArgPheLybValArgMeGlyGly 20
DB 1 ATGGCTCATTCAGAACGCGTCTAAAGAAAGAAATGATCAATGAATATACCAATGAGGAG 60
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_p2n model

Run on: January 12, 2006, 10:08:26 ; Search time 793 seconds  
(without alignments)  
2346.293 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVYKEFMFKVMEG.....EDYTVIEQYERTEGRHHLFL 225

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues  
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODL=Frame+P2n.model -DEV=xlp  
-Q=/cgn2\_1/USPFO.epool/p/US10006922/runat\_10012006.162409.18977/app\_query.fasta\_1.391  
-DB=Published.Applications\_NA\_Main -QFMT=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1  
-MATRIX=bloum62 -TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext  
-HEA\_SIZE=500 -MUTLEN=0  
-USER=US10006922@cgn2\_1.1549@runat\_10012006.162409.18977 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA\_Main:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1214	100.0	678	3 US-09-967-772-6
2	1214	100.0	678	5 US-10-006-922-11
3	1214	100.0	678	5 US-10-081-864-7
4	1214	100.0	678	5 US-10-121-258-2
5	1214	100.0	678	6 US-10-315-920-1
6	1214	100.0	678	6 US-10-132-067-3
7	1214	100.0	678	6 US-10-335-517-6

8	1214	100.0	678	6 US-10-334-268-6	Sequence 6, Appl1
9	1214	100.0	678	7 US-10-311-030-5	Sequence 5, Appl1
10	1214	100.0	678	8 US-10-656-029-21	Sequence 21, Appl1
11	1214	100.0	678	9 US-10-505-486-27	Sequence 27, Appl1
12	1214	100.0	678	9 US-10-844-064A-1	Sequence 1, Appl1
13	1214	100.0	678	9 US-10-931-304-2	Sequence 2, Appl1
14	1214	100.0	859	3 US-09-999-745-66	Sequence 66, Appl1
15	1214	100.0	859	3 US-09-866-538-11	Sequence 11, Appl1
16	1214	100.0	859	3 US-09-794-308-11	Sequence 11, Appl1
17	1214	100.0	859	3 US-09-865-291-11	Sequence 11, Appl1
18	1214	100.0	859	3 US-10-433-640-12	Sequence 12, Appl1
19	1214	100.0	859	8 US-10-885-988-11	Sequence 11, Appl1
20	1214	100.0	859	8 US-10-857-632-11	Sequence 11, Appl1
21	1214	100.0	3311	3 US-09-797-496B-3	Sequence 3, Appl1
22	1210	99.7	681	5 US-10-006-922-35	Sequence 35, Appl1
23	1210	99.7	681	5 US-10-121-258-3	Sequence 3, Appl1
24	1210	99.7	681	5 US-10-121-258-23	Sequence 23, Appl1
25	1210	99.7	681	7 US-10-311-030-8	Sequence 8, Appl1
26	1210	99.7	681	9 US-10-931-304-3	Sequence 3, Appl1
27	1210	99.7	681	9 US-10-931-304-23	Sequence 23, Appl1
28	1210	99.7	713	7 US-10-311-030-11	Sequence 11, Appl1
29	1210	99.7	713	7 US-10-311-030-12	Sequence 12, Appl1
30	1210	99.7	723	5 US-10-152-296-1	Sequence 1, Appl1
31	1210	99.7	723	7 US-10-739-656-1	Sequence 1, Appl1
32	1210	99.7	1638	5 US-10-214-932-51	Sequence 51, Appl1
33	1210	99.7	1647	5 US-10-214-932-75	Sequence 75, Appl1
34	1210	99.7	4692	6 US-10-161-403-29	Sequence 29, Appl1
35	1210	99.7	4692	7 US-10-433-640-16	Sequence 16, Appl1
36	1210	99.7	4692	10 US-11-006-076-29	Sequence 29, Appl1
37	1210	99.7	5436	9 US-10-169-050-46	Sequence 46, Appl1
38	1210	99.7	6984	5 US-10-001-189-45	Sequence 45, Appl1
39	1210	99.7	7910	9 US-10-169-050-20	Sequence 20, Appl1
40	1210	99.7	9320	8 US-10-471-065-20	Sequence 20, Appl1
41	1210	99.7	9658	8 US-10-609-019-4	Sequence 4, Appl1
42	1210	99.7	9678	8 US-10-609-019-3	Sequence 3, Appl1
43	1210	99.7	10263	8 US-10-609-019-2	Sequence 2, Appl1
44	1207	99.4	678	6 US-10-315-920-3	Sequence 3, Appl1
45	1203	99.1	7508	7 US-10-742-828-4	Sequence 4, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-967-772-6  
/ Sequence 6, Application US/09967772  
/ Patent No. US20020164577A1  
/ GENERAL INFORMATION:  
/ APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
/ APPLICANT: TSIEN, Roger  
/ APPLICANT: GONZALEZ, Jesus  
/ TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
/ FILE REFERENCE: REGEN1290-5  
/ CURRENT APPLICATION NUMBER: US/09/967,772  
/ CURRENT FILING DATE: 2001-09-28  
/ PRIOR APPLICATION NUMBER: US 09/459,956  
/ PRIOR FILING DATE: 1999-12-13  
/ PRIOR APPLICATION NUMBER: US 08/765,860  
/ PRIOR FILING DATE: 1996-12-19  
/ PRIOR APPLICATION NUMBER: PCT/ US96/09652  
/ PRIOR FILING DATE: 1996-06-06  
/ PRIOR APPLICATION NUMBER: US 08/481,977  
/ PRIOR FILING DATE: 1995-06-07  
/ NUMBER OF SEQ ID NOS: 22  
/ SOFTWARE: PatentIn version 3.0  
/ SEQ ID NO 6  
/ LENGTH: 678  
/ TYPE: DNA  
/ ORGANISM: Discosoma sp "red"  
/ US-09-967-772-6  
Alignment Scores: 6.83e-143 Length: 678  
Pred. No.: 1214.00 Matches: 225  
Score:

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-09-967-772-6 (1-678)

QY 1 MetArgSerSerLybAanValIleLybGluPhneMetArgPheLybValArgMetGluGly 20
   1 ATGAGGCTCTCCAGAAATGTTATCAAGAGTTCTAGAGTTTAAGTTCCGATCGAAGGA 60
   |||||
QY 21 ThrValAsnGlyYhiSgluPheGluIleGluGlyGluGlyArgProTyrgLugly 40
   61 ACGGTCAATGGCGACGAGTTGAATAGAAAGCGAAGAGAGGGAGGCCATAGAAAGGC 120
   |||||
QY 41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTPAspIle 60
   121 CACAAATACCGTAAAGCTTAAAGTAACCAAGGGGAGACCTTGGCAATTTGCTTGGATATT 180
   |||||
QY 61 LeuSerProGlnPheGlnTyrgLySerLybValTyrgValLybShiSProAlaAspIlePro 80
   181 TTGTCACCAACATTTCAAGTATGAAAGCAAGGTATATGTCAAGCAACCTGCCGACATACCA 240
   |||||
QY 81 AspTyrgLybLeuSerPheProGluGlyPheLybTPGluArgValMetAsnPheGlu 100
   241 GACTATATAAAGCTGTCATTTCTGAAAGATTTAAATGGGAAAGGTCATGAATTTGAA 300
   |||||
QY 101 AspGlyGlyValIleThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyrg 120
   301 GACGCTGGCGCGTCTTACTGTAAACCAAGATTCAGATTGCGAGATGGCTGTTTCATCTAC 360
   |||||
QY 121 LybValLybPheIleGlyValIleAsnPheProSerAspGlyProValMetGlnLybLeuThr 140
   361 AAGGTCAAGTTCAATGGCGTAACTTTCTTCCATGAGACCTGTTATGCAAAAGAAAGCA 420
   |||||
QY 141 MetGlyTPGluIleAspSerThrGluArgLeuTyrgProArgAspGlyValLeuLybGly 160
   421 ATGGGCTGGGAAGCCAGACAGTGAAGCTTTGTATCTCGATGAGCGTGGTGAAGAGAG 480
   |||||
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyYhiSlyrLeuValGluPheLybSerIle 180
   481 ATTCATTAAGGCTCTGAAGCTGAAGACGGTGTCAATTAACCTAGATTCAAAAGATATT 540
   |||||
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrgTyrgValIleAspSerLybLeuAsp 200
   541 TACATGGCAAGAAAGCTGTGACACTACAGGGTACTACTATGTGACTCCAAACTGGAT 600
   |||||
QY 201 IleThrSerHisAsnGluAspTyrgThrIleValGluGlnTyrgIuArgThrGluGlyArg 220
   601 ATAACAAGCCACAAGAACTATACATCGTTGAGCAGTATGAAAGAACGAGGAGCGC 660
   |||||
QY 221 HisHisLeuPheLeu 225
   661 CACCATCTGTTCTT 675
   |||||
DB 661 CACCATCTGTTCTT 675
   |||||

RESULT 2
US-10-006-922-11
; Sequence 11, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A
; APPLICANT: Fredkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tetsikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1 Chromophores/Fluorophores and
; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
```

```

; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma species
US-10-006-922-11

Alignment Scores:
Pred. No.: 6,83e-143 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-006-922-11 (1-678)

QY 1 MetArgSerSerLybAanValIleLybGluPhneMetArgPheLybValArgMetGluGly 20
   1 ATGAGGCTCTCCAGAAATGTTATCAAGAGTTTATGAGTTTAAGTTCCGATCGAAGGA 60
   |||||
QY 21 ThrValAsnGlyYhiSgluPheGluIleGluGlyGluGlyArgProTyrgLugly 40
   61 ACGGTCAATGGCGACGAGTTGAATAGAAAGCGAAGAGAGGGAGGCCATAGAAAGGC 120
   |||||
QY 41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTPAspIle 60
   121 CACAAATACCGTAAAGCTTAAAGTAACCAAGGGGAGACCTTGGCAATTTGCTTGGATATT 180
   |||||
QY 61 LeuSerProGlnPheGlnTyrgLySerLybValTyrgValLybShiSProAlaAspIlePro 80
   181 TTGTCACCAACATTTCAAGTATGAAAGCAAGGTATATGTCAAGCAACCTGCCGACATACCA 240
   |||||
QY 81 AspTyrgLybLeuSerPheProGluGlyPheLybTPGluArgValMetAsnPheGlu 100
   241 GACTATATAAAGCTGTCATTTCTGAAAGATTTAAATGGGAAAGGTCATGAATTTGAA 300
   |||||
QY 101 AspGlyGlyValIleThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyrg 120
   301 GACGCTGGCGCGTCTTACTGTAAACCAAGATTCAGATTGCGAGATGGCTGTTTCATCTAC 360
   |||||
QY 121 LybValLybPheIleGlyValIleAsnPheProSerAspGlyProValMetGlnLybLeuThr 140
   361 AAGGTCAAGTTCAATGGCGTGAAGCTTTCTTCCGATGAGACTGTATTGCAAAAGAAAGCA 420
   |||||
QY 141 MetGlyTPGluIleAspSerThrGluArgLeuTyrgProArgAspGlyValLeuLybGly 160
   421 ATGGGCTGGGAAGCCAGCACTGAGCCTTTGTATCTCTGATGAGCGTGGTGAAGAGAG 480
   |||||
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyYhiSlyrLeuValGluPheLybSerIle 180
   481 ATTCATTAAGGCTCTGAAGCTGAAGAAAGCGGTGTCAATTAAGTGAATTCAAAAGATATT 540
   |||||
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrgTyrgValIleAspSerLybLeuAsp 200
   541 TACATGGCAAGAAAGCTGTGACACTACAGGGTACTACTATGTTCATCTCCAAACTGGAT 600
   |||||
QY 201 IleThrSerHisAsnGluAspTyrgThrIleValGluGlnTyrgIuArgThrGluGlyArg 220
   601 ATAACAAGCCACAAGAACTATACATCGTTGAGCAGTATGAAAGAACGAGGAGCGC 660
   |||||
QY 221 HisHisLeuPheLeu 225
   |||||
```

```

Db          661 CACCATCTGTTCTT 675

RESULT 3
US-10-081-864-7
; Sequence 7, Application US/10081864
; Publication No. US20030022287A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Yanushevich, Yuriy
; APPLICANT: Savitsky, Alexander
; APPLICANT: Fradkov, Arcady
; TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: CLON-067
; CURRENT APPLICATION NUMBER: US/10/081,864
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: 10/006,922
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/270,983
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Drosophila sp
US-10-081-864-7

```

Alignment Scores:		6,836-143	Length:	678
Pred. No.:	Score:	1214.00	Matches:	225
Percent Similarity:		100.00%	Conservative:	0
Best Local Similarity:		100.00%	Mismatches:	0
Query Match:		100.00%	Indels:	0
DB:		5	Gaps:	0
US-10-006-922A-12 (1-225) x US-10-081-864-7 (1-678)				
QY	1	MeCaRgSeSerSeLyAAsnValIleLyegInuPhMeCaRtPhLyVaIArgMeGInuLy	20	
Db	1	ATGCCCTCTCCAAAGACGTCTCAAGAGTTCAATGCCCTTCAGGTCCGATGAGAGGC	60	
QY	21	ThVaLaAnGjYHLeGInuPhGInuIleGInuLyGInuLyGInuLyAArgProTyrGInuLy	40	
Db	61	ACCGTGAACGGCCACGAGTTCAGATTCGAGGCGCGAGGCGAGAGGCGCGCCCTCAACAGGGC	120	
QY	41	HiSaEnThThVaIlyLeuLyVaIThrLySgLyGjYProLeuProPhaIaTrSpaIle	60	
Db	121	CACCAACACCGTAMCTGAAGGTGACCAAGGGCGGGCCCTCGCTTCGCTGGAGACATC	180	
QY	61	LeuSerProGInuPhGInuTyrGjYSerLyVaIYrVaIlyHiSaProLaAapIlePro	80	
Db	181	CTGTCCCGCCAGTTCCAGTACGGGCTCAAGGTACGTAGACACCCCGCCGACATCCCC	240	
QY	81	AaPyTyrLyLyLeuSerPheProGInuLyPheLySTrPGLuArgVaIMetAsnPhGInu	100	
Db	241	GACTCAACMAACCTGCTCTCCCGAGGGCTTCMAETGGAGACGGGTATGAACCTTCAG	300	
QY	101	AaPyLyGjYVaIyVaIThrVaIThrGInaESeSerLeuGInaAaPyLYCyAaPheIleYr	120	
Db	301	GACGGCGGCGGTGTGACCTGACCCAGACACTCTCCCTTCAGAGACGGGTCTTCAATC	360	
QY	121	LyVaIlyAaPheIleGjYVaIaAnPheProSeAaPyLYProVaIMeGInuLySjYTr	140	
Db	361	AAGGTGAAGTTCACTCGCGCTGAACCTTCCCTCCGACGGCGCGGTATGCAAGAGAACCC	420	
QY	141	MeGjYTrPGLuAaSeThrGjYArgLeuYrProArgAaPyLYaIleuLySgLyGInu	160	
Db	421	ATGGGCTGGAGGCTTCACCGAGCGCTGTACCCCGGACGGGCTCTGAAAGGCGAG	480	
QY	161	IlEnaLyAaLeuLyLeuLyAaPyLYGjYHieTyrLeuVaIcInuPhLySerIle	180	

```

Db      301 GACGGTGGCGTCTTACTGTAAACCAAGATTCCAGTTTGAGAGATGGCTGTTTCAATC 360
Qy      121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr 140
Db      361 AAGGTCAAGTTCATGTGGGTGAATCTTCTCCATGAGCTGTATTCAGAAAGAGCA 420
Qy      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGly 160
Db      421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCTCTGATGGCGATGGATGAGAG 480
Qy      161 ILeHsLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSer 180
Db      481 ATTCATAAGCGCTCGAGACTGAAAGACGGTGTGATTACTAGTTGAATTCAGAAAG 540
Qy      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeu 200
Db      541 TACATGGCAAGAAAGCCGTGTGACGTAACAGGGTACTACTATGTTGACTCCAAAC 600
Qy      201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGly 220
Db      601 ATACCAAGCCCAAGACGAGACTATACATCGTTGACGATGAAAGAACCGAGGAGC 660
Qy      221 HisHisLeuPheLeu 225
Db      661 CACCATCTGTTCTT 675

```

## RESULT 5

```

US-10-315-920-1
; Sequence 1, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Terzikikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant of sequence from Discoosoma sp.
US-10-315-920-1

```

## Alignment Scores:

```

Pred. No.: 6,83e-143      Length: 678
Score: 1214.00      Matches: 225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 6      Gaps: 0

```

US-10-006-922a-12 (1-225) x US-10-315-920-1 (1-678)

```

Qy      1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db      1 ATGGGCTCTCCCAAGACGTCAATCAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGAGGC 60
Qy      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGluGlyArgProTyrGluGly 40
Db      61 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGAGGGCGAGGGCCGCGCTTACGAGGCG 120
Qy      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

```

```

Db      121 CACAACACCGTGAAGCTGAAGTGAACCAAGGGCGGCCCTTCGCTTCGCTGGGAGATC 180
Qy      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisPProAlaAspIlePro 80
Db      181 CTGTCCTCCCAAGTTCAGTACGGCTCCAAAGTGTACGTGAAGCACCCCGCGGACATCCCC 240
Qy      81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrPglLysArgValMetAsnPheGlu 100
Db      241 GACTACAAAGAACGTCTCTTCCCGAGGGCTTCAGAGTGGAGGCCCGGATGAACTTGAG 300
Qy      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      301 GACGCGCGCGTGGTGAACCGTGAACCAAGACTCTCTCGACGAGCGGCTTCATCTAC 360
Qy      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db      361 AAGGTGAAGTTCATCGCGGTGAATCTTCCCTCCGACGGCCCGGTGAAGCAAGAAACCC 420
Qy      141 MetGlyTyrPglLysSerThrGluArgLeuTyrProArgAspGlyValLeuLysGly 160
Db      421 ATGGGCTGGGAAGGCTCTCACCGAGCGCTGTACCCCGCAAGCGCGTGTGAAGGGCAG 480
Qy      161 ILeHsLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSer 180
Db      481 ATCCACAAGCCCTCGAAGCTGAAGGACGCGGCACTACTGTGGAGTTCAAGTCCATC 540
Qy      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeu 200
Db      541 TACATGGCAAGAAAGCCGTGTGACGCTGCCGCTACTACTACGTGAGCTCCAAAGCTG 600
Qy      201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGly 220
Db      601 ATACCTCCCAAGACGAGACTACACATGTGAGAGATACGAGCGCAAGAGGGCGCG 660
Qy      221 HisHisLeuPheLeu 225
Db      661 CACCACCTGTTCTT 675

```

## RESULT 6

```

US-10-132-067-3
; Sequence 3, Application US/10132067
; Publication No. US20030203355A1
; GENERAL INFORMATION:
; APPLICANT: Bradbury, Andrew
; APPLICANT: Zeytun, Ahmet
; APPLICANT: Waldo, Geoffrey
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Fluorobodies: Binding ligands with intrinsic
; FILE REFERENCE: 021362-000600US
; CURRENT APPLICATION NUMBER: US/10/132,067
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discoosoma sp.
; FEATURE:
; OTHER INFORMATION: red fluorescent protein (dRED)
; NAME/KEY: CDS
; LOCATION: (1)-(678)
; OTHER INFORMATION: dRED
US-10-132-067-3

```

## Alignment Scores:

```

Pred. No.: 6,83e-143      Length: 678
Score: 1214.00      Matches: 225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 6      Gaps: 0

```

US-10-006-922a-12 (1-225) x US-10-132-067-3 (1-678)

QY 1 MetArgSerSerIyEAAnValIIlelyeGIuPhMeLArGpHeLyValAArgMeGluGIy 20  
 Db 1 ATGAGGCTTCTCCAGAAAGTTATCAAGAGTTCAAGAGTTTAAAGTTCCGATGGAGAGA 60

QY 21 ThrValAenGIyHISgluPhEgluIIeGIuGIyGIuGIyGIuGIyArgProTyrgIuGIy 40  
 Db 61 ACGGTCATGGGCAAGCTTTGAATTAAGAGCGAAGAGAGGGGAGGCCATACGAAGGC 120

QY 41 HIsAenThrValIyLeuLyEValIThrIySglYgIyProLeuProPhaIATrPAspIle 60  
 Db 121 CACATATCCGTAAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGGCTTGGGATATT 180

QY 61 LeuSerProGInPhEgInTyrgIySerIyValIyTrValIyVHISProAlAspIlePro 80  
 Db 181 TTGTCAACAATTCAGATATGAAAGCAAGGTATATGTCAAGCACCTGGCGACATACCA 240

QY 81 AspTyrlYsLyLeuSerPheProGIuGIyPheLySTrPGIuArgValMeLAsnPhEglu 100  
 Db 241 GACCTATAAAGCTGTCTATTCCTGAAGATTAAATGGAAAGGTCATGAACCTTTGA 300

QY 101 AspGIyGIyValIyThrValIThrGInAsPserSerLeuGInAspGIyCyPHeIleTy 120  
 Db 301 GACGGTGGCGTGTACTGTAAACCCAGATTCAGATTGGCAGATGGCTGTTCATCTAC 360

QY 121 LySValIyVpHeIIeGIyValIAsnPhEProSerAspGIyProValMeGInLySlySThr 140  
 Db 361 AAGGTCAAGTTCAATGGCGTGAACCTTCTCCGATGACCTGTATATGCAAAAGAGACA 420

QY 141 MetGIyTrPGIuAlAsEThrGIuArgLeuTyProArgAspGIyValIleuLySglYGIu 160  
 Db 421 ATGGCTGGGAAGCCAGCACTGAGCTTTGTATCTCCGTGATGGCGTGTGAAGAGAG 480

QY 161 ILeHISlyValIleuLyLeuLyAspGIyGIyHISlyrLeuValIGluPhElySserIle 180  
 Db 481 ATTCAATTAAGGCTCTGAAGCTGAAGACGGTGTCTACTTACCTAGTGAATTCAAAAGTATT 540

QY 181 TyrMeLAlLyLyVpProValIGInLeuProGIyTyTrTyTrValIAsPserLySleuAsP 200  
 Db 541 TACATGGCAAAAGCTGTGACGTACCAAGGTACTACTATGTTGACTCCAACTGGAT 600

QY 201 ILeThrSerHISAsnGIuAspTyTrThrIleValIGluGInTyrgIuArgThrgIuGIyArg 220  
 Db 601 ATTAACAAGCAACAAGACATTAACAATCTTTGACGATATGAAGAAGAACCGAGGGAGCC 660

QY 221 HIsHISleuPhELeu 225  
 Db 661 CACCATCTGTCTCTT 675

RESULT 7  
 US-10-335-517-6  
 ; Sequence 6, Application US/10335517  
 ; Publication No. US20030207248A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 ; APPLICANT: TSIEH, Roger  
 ; APPLICANT: GONZALEZ, Jesus  
 ; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
 ; FILE REFERENCE: REGEN1290-5  
 ; CURRENT APPLICATION NUMBER: US/10/335,517  
 ; PRIORITY FILING DATE: 2002-12-31  
 ; PRIOR APPLICATION NUMBER: US/09/967,772  
 ; PRIOR FILING DATE: 2001-09-28  
 ; PRIOR APPLICATION NUMBER: US 09/459,956  
 ; PRIOR FILING DATE: 1999-12-13  
 ; PRIOR APPLICATION NUMBER: US 08/765,860  
 ; PRIOR FILING DATE: 1996-12-19  
 ; PRIOR APPLICATION NUMBER: PCT/ US96/09652  
 ; PRIOR FILING DATE: 1996-06-06  
 ; PRIOR APPLICATION NUMBER: US 08/481,977  
 ; PRIOR FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO: 6  
 ; LENGTH: 678  
 ; TYPE: DNA  
 ; ORGANISM: *Dicosoma sp "red"*  
 US-10-335-517-6

Alignment Scores:  
 Pred. No.: 6,83e-143 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-335-517-6 (1-678)

QY 1 MetArgSerSerIyEAAnValIIlelyeGIuPhMeLArGpHeLyValAArgMeGluGIy 20  
 Db 1 ATGAGGCTTCTCCAGAAAGTTATCAAGAGTTCAAGAGTTTAAAGTTCCGATGGAGAGA 60

QY 21 ThrValAenGIyHISgluPhEgluIIeGIuGIyGIuGIyGIuGIyArgProTyrgIuGIy 40  
 Db 61 ACGGTCATGGGCAAGCTTTGAATTAAGAGCGAAGAGAGGGGAGGCCATACGAAGGC 120

QY 41 HIsAenThrValIyLeuLyEValIThrIySglYgIyProLeuProPhaIATrPAspIle 60  
 Db 61 CACATATCCGTAAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGGCTTGGGATATT 180

QY 61 LeuSerProGInPhEgInTyrgIySerIyValIyTrValIyVHISProAlAspIlePro 80  
 Db 181 TTGTCAACAATTCAGATATGAAAGCAAGGTATATGTCAAGCACCTGGCGACATACCA 240

QY 81 AspTyrlYsLyLeuSerPheProGIuGIyPheLySTrPGIuArgValMeLAsnPhEglu 100  
 Db 241 GACCTATAAAGCTGTCTATTCCTGAAGATTAAATGGAAAGGTCATGAACCTTTGA 300

QY 101 AspGIyGIyValIyThrValIThrGInAsPserSerLeuGInAspGIyCyPHeIleTy 120  
 Db 301 GACGGTGGCGTGTACTGTAAACCCAGATTCAGATTGGCAGATGGCTGTTCATCTAC 360

QY 121 LySValIyVpHeIIeGIyValIAsnPhEProSerAspGIyProValMeGInLySlySThr 140  
 Db 361 AAGGTCAAGTTCAATGGCGTGAACCTTCTCCGATGACCTGTATATGCAAAAGAGACA 420

QY 141 MetGIyTrPGIuAlAsEThrGIuArgLeuTyProArgAspGIyValIleuLySglYGIu 160  
 Db 421 ATGGCTGGGAAGCCAGCACTGAGCTTTGTATCTCCGTGATGGCGTGTGAAGAGAG 480

QY 161 ILeHISlyValIleuLyLeuLyAspGIyGIyHISlyrLeuValIGluPhElySserIle 180  
 Db 481 ATTCAATTAAGGCTCTGAAGCTGAAGACGGTGTCTACTTACCTAGTGAATTCAAAAGTATT 540

QY 181 TyrMeLAlLyLyVpProValIGInLeuProGIyTyTrTyTrValIAsPserLySleuAsP 200  
 Db 541 TACATGGCAAAAGCTGTGACGTACCAAGGTACTACTATGTTGACTCCAACTGGAT 600

QY 201 ILeThrSerHISAsnGIuAspTyTrThrIleValIGluGInTyrgIuArgThrgIuGIyArg 220  
 Db 601 ATTAACAAGCAACAAGACATTAACAATCTTTGACGATATGAAGAAGAACCGAGGGAGCC 660

QY 221 HIsHISleuPhELeu 225  
 Db 661 CACCATCTGTCTCTT 675

RESULT 8  
 US-10-334-288-6  
 ; Sequence 6, Application US/10334288  
 ; Publication No. US20040002123A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 ; APPLICANT: TSIEH, Roger

APPLICANT: GONZALEZ, Jesus  
TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
FILE REFERENCE: RESEN1290-5  
CURRENT FILING DATE: 2002-12-31  
PRIOR APPLICATION NUMBER: US/10/334,288  
PRIOR FILING DATE: 2002-12-31  
PRIOR APPLICATION NUMBER: US/09/967,772  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: US 09/459,956  
PRIOR FILING DATE: 1999-12-13  
PRIOR APPLICATION NUMBER: US 08/765,860  
PRIOR FILING DATE: 1996-12-19  
PRIOR APPLICATION NUMBER: PCT/ US96/09652  
PRIOR FILING DATE: 1996-06-06  
PRIOR APPLICATION NUMBER: US 08/481,977  
PRIOR FILING DATE: 1995-06-07  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 6  
LENGTH: 678  
TYPE: DNA  
ORGANISM: *Discoosoma* sp "red"  
US-10-334-288-6

Alignment Scores:  
Pred. No.: 6,83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-334-288-6 (1-678)

Qy 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20  
Db 1 ATGAGGCTTCCAAAGATGTATCAAGAGTTCAATGAGTTTAAAGTTCCGATGGAAGA 60  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyGluGly 40  
Db 61 ACGGTCATGGGACACGATTTGAATAGAGCGAAGAGGGGAGGCCATTCAGAAAGGC 120  
Qy 41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAAATACCGTAAAGCTTAAAGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180  
Qy 61 LeuSerProGlnPheGlnTyGlySerLySValTyValLySHisProAlaAspIlePro 80  
Db 181 TTGTACACCAACAATTCAGTATGGAAGCAAGGTATATGCAAGCACCTCGCCGACATACCA 240  
Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATATAAAGCTGTCTCTTCCTGAGAGATTTAAATGGGAAAGGTCATGAACTTTGAA 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerLeuGlnAspGlyCySPhelIleTyr 120  
Db 301 GAGCGTGGCGTCTGTAATCGTAACCAAGATTCCAGTTTGACAGATGGCTGTTTCATCTAC 360  
Qy 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
Db 361 AAGGTCAAGTTCAATTTGGGTGAACTTTCTTCCTGATGACCTGTTAAGCAAAAGACAA 420  
Qy 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyProArgAspGlyValLeuLySGLyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCTGTAGGCGTGTGAAAGAGAG 480  
Qy 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyLeuValGluPheLySLeuThr 180  
Db 481 ATTCAATAGGCTCGAAGCTGAAAGACGGTGTCAATTACCAAGTTGAATCAAAAGATATT 540  
Qy 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyTyrValAspSerLySLeuAsp 200  
Db 541 TACATGGCAAAAGACCTGTGCACTACAGGGGTATCTATGTGTGACTCCAAACCTGGAT 600

Qy 201 IleHisSerHisAsnGluAspTyrThrIleValGluGlnTyGluArgThrGluGlyArg 220  
Db 601 ATACCAAGCCCAACCAAGACCTATACATCTTGAGCAGTATGAAAGACGAGGAGCGC 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCTT 675

## RESULT 9

US-10-311-030-5  
Sequence 5, Application US/10311030  
Publication No. US20040171107A1  
GENERAL INFORMATION:  
APPLICANT: Nelson, David  
APPLICANT: Zamiatara, Elize  
APPLICANT: Trien, Roger  
TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS  
FILE REFERENCE: 15916-032US1  
CURRENT APPLICATION NUMBER: US/10/311,030  
CURRENT FILING DATE: 2002-12-10  
PRIOR APPLICATION NUMBER: PCT/US01/04625  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: US 60/184,732  
PRIOR FILING DATE: 2000-02-23  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 678  
TYPE: DNA  
ORGANISM: *Discoosoma* sp.  
US-10-311-030-5

Alignment Scores:  
Pred. No.: 6,83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-311-030-5 (1-678)

Qy 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLySValArgMetGluGly 20  
Db 1 ATGAGGCTTCCAAAGATGTATCAAGAGTTCAATGAGTTTAAAGTTCCGATGGAAGA 60  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyGluGly 40  
Db 61 ACGGTCATGGGACACGATTTGAATAGAGCGAAGAGGGGAGGCCATTCAGAAAGGC 120  
Qy 41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAAATACCGTAAAGCTTAAAGTAACCAAGGGGGACCTTTGCCATTTGCTGGGATATT 180  
Qy 61 LeuSerProGlnPheGlnTyGlySerLySValTyValLySHisProAlaAspIlePro 80  
Db 181 TTGTACACCAACAATTCAGTATGGAAGCAAGGTATATGCAAGCACCTCGCCGACATACCA 240  
Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATATAAAGCTGTCTCTTCCTGAGAGATTTAAATGGGAAAGGTCATGAACTTTGAA 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerLeuGlnAspGlyCySPhelIleTyr 120  
Db 301 GAGCGTGGCGTCTGTAATCGTAACCAAGATTCCAGTTTGACAGATGGCTGTTTCATCTAC 360  
Qy 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
Db 361 AAGGTCAAGTTCAATTTGGGTGAACTTTCTTCCTGATGACCTGTTAAGCAAAAGACAA 420  
Qy 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyProArgAspGlyValLeuLySGLyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCTGTAGGCGTGTGAAAGAGAG 480



QY 161 ILEHSLVALALEULYLEULYSAAPGLYGLYHISLYRLEUVALGLUPHELYSERILE 180  
DB 481 ATTCTAAGGCTCTGACACTGAAAGACGGTGCTATTCTAGTTGAATTCAAAAGTATT 540  
QY 181 TYRMECALALYSLYSPROVALGLNLEUPROGLYTYRTRYRVALAASPSELYLEUASP 200  
DB 541 TACATGGCAAGAACCTGTGACGCTACAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 ILEHTRSERHISAANGUASPTYRTHRIEVALIGUGINTYRGUARGTHGLUGLYARG 220  
DB 601 ATACCAACCCCAACGAAGACTATACATCGTTGACGATGAAAGAACGAGGAGCC 660  
QY 221 HISLSELPHELEU 225  
DB 661 CACCATCTGTTCTT 675  
RESULT 10  
US-10-656-029-21  
; Sequence 21, Application US/10656029  
; Publication No. US2005000367A1  
; GENERAL INFORMATION:  
; APPLICANT: VERTEX PHARMACEUTICALS INC.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF  
; FILE REFERENCE: VPI/02-143WO2  
; CURRENT APPLICATION NUMBER: US/10/656,029  
; PRIOR FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: 60/408,297  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 21  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Drosophila sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(678)  
; OTHER INFORMATION: Fluorescent protein  
US-10-656-029-21  
Alignment Scores:  
Pred. No.: 6,83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-006-922a-12 (1-225) x US-10-656-029-21 (1-678)  
QY 1 MetAAGSerSerLyAsnValIleLySGlUpHMeCArgPhelyValAArgMeGluGly 20  
DB 1 ATGAGGTCCTCCAAAGATGTATCAAGAGTTCAATGAGTTTAAAGTTGCAATGGAAGCA 60  
QY 21 ThrValAAsnGLYHISGLUPheGluIleGluGLYGLUGLYARGPProTYRGLUGLY 40  
DB 61 ACGGTCATAGGCAAGATTGAAATAGAAAGCGAAGAGGAGGAGGCCATACGAAGC 120  
QY 41 HISAsnThrValLyLeuLyValThrLySGLYGLYProLeuProPheAATPAPile 60  
DB 121 CACAATACCGTAAGCTTAAGTTAACCAAGGGGAGCCTTGGCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTYRGLYSerLyValTYRValLyHISProAAsPilePro 80  
DB 181 TTGTCAACCAATTTAGTAGAAGCAAGTATATGTCACACCTTCCGACATACCA 240  
QY 81 AAsPTYRlyLyLeuSerPheProGluGLYPhelySerTrpGluArgValMetAsnPheliu 100  
DB 241 GACTATTAAGAGTGTGCTTCTGAAAGATTTAAATGGAAAGGCTCATGAATTTGA 300  
QY 101 AAsGlyGLYValValThrValThrGlnAAsSerSerLeuGlnAAsGlyCyAsPheIleTYR 120

DB 301 GACGATGCGCTGTTACTGTACCAAGATTCAGTTTGCGAGATGGCTGTTTCATCTAC 360  
QY 121 LyValLySPheIleGlyValAAsnPheProSerAAsGlyProValMetGlnLySThr 140  
DB 361 AAGGTCAAGTTCATTTGGGTGAACCTTCTCCATGAGACCTGTATATCAAAAGACACA 420  
QY 141 MetGLYTPGLUALaserThrGluArgLeuTYRProArgAAsGlyValIleLyGLYGLU 160  
DB 421 ATGGCTGGGAAGCAGACACTGAGCGTTTGATCTCGTGAATGCGGTGTAAGAGAGAG 480  
QY 161 ILEHSLVALALEULYLEULYSAAPGLYGLYHISLYRLEUVALGLUPHELYSERILE 180  
DB 481 ATTCTAAGGCTCTGACACTGAAAGACGGTGCTATTCTAGTTGAATTCAAAAGTATT 540  
QY 181 TYRMECALALYSLYSPROVALGLNLEUPROGLYTYRTRYRVALAASPSELYLEUASP 200  
DB 541 TACATGGCAAGAACCTGTGACGCTACAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 ILEHTRSERHISAANGUASPTYRTHRIEVALIGUGINTYRGUARGTHGLUGLYARG 220  
DB 601 ATACCAACCCCAACGAAGACTATACATCGTTGACGATGAAAGAACGAGGAGCC 660  
QY 221 HISLSELPHELEU 225  
DB 661 CACCATCTGTTCTT 675  
RESULT 11  
US-10-505-486-27  
; Sequence 27, Application US/10505486  
; Publication No. US20050118639A1  
; GENERAL INFORMATION:  
; APPLICANT: Takeda Chemical Industries, Ltd.  
; TITLE OF INVENTION: Determination of a ligand  
; FILE REFERENCE: P03-0006PCT  
; CURRENT APPLICATION NUMBER: US/10/505,486  
; PRIOR FILING DATE: 2004-08-20  
; PRIOR APPLICATION NUMBER: JP 2002-45728  
; PRIOR FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: JP 2002-213949  
; PRIOR FILING DATE: 2002-07-23  
; PRIOR APPLICATION NUMBER: JP 2002-298237  
; NUMBER OF SEQ ID NOS: 233  
; SEQ ID NO 27  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Drosophila sp.  
US-10-505-486-27  
Alignment Scores:  
Pred. No.: 6,83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-10-006-922a-12 (1-225) x US-10-505-486-27 (1-678)  
QY 1 MetAAGSerSerLyAsnValIleLySGlUpHMeCArgPhelyValAArgMeGluGly 20  
DB 1 ATGAGGTCCTCCAAAGATGTATCAAGAGTTCAATGAGTTTAAAGTTGCAATGGAAGCA 60  
QY 21 ThrValAAsnGLYHISGLUPheGluIleGluGLYGLUGLYARGPProTYRGLUGLY 40  
DB 61 ACGGTCATAGGCAAGATTGAAATAGAAAGCGAAGAGGAGGAGGCCATACGAAGC 120  
QY 41 HISAsnThrValLyLeuLyValThrLySGLYGLYProLeuProPheAATPAPile 60  
DB 121 CACAATACCGTAAGCTTAAGTTAACCAAGGGGAGCCTTGGCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTYRGLYSerLyValTYRValLyHISProAAsPilePro 80

Db 181 TTGTCACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACTCCGCGCATATCA 240

Qy 81 AepTyrIyLysLeuSerPheProGluGlyPheIyTfPgluaTgValMetAsnPheGlu 100

Db 241 GACTATATAAAAGCTGTCTTCCTGAAAGATTAAATGGGAAAGGCTCATCACTTTGAA 300

Qy 101 AaspGlyGlyValIValThrValThrGlnAspSerSerLeuGlnaAspGlyCyAspPheIleTyr 120

Db 301 GACGGTGGCGTGTACTGTAAACCCAGATTCCAGTTGCAAGATGGCTGTTCATCTAC 360

Qy 121 LysValIySphelIegIyValaAsnPheProSerAspGlyProValMetGlnLysLeuThr 140

Db 361 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCAGATGACCTGTTATGCAAAAGAGCA 420

Qy 141 MetGlyTfPgluaIAserThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 421 ATGGGCTGGGAAGCAGACACTGAGCGTTTGTATCTCTCGATGGCGTGTGAAAGGAGAG 480

Qy 161 ILehIySphalaleuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 481 ATTCAATAAAGCTGTGAAGCTAAAGACGGTGTCTATTAAGTTGAATTCAAAAGATTT 540

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValaAspSerLysLeuAsp 200

Db 541 TACATGGCAAGAAAGCCGTGTGACGTAACGAGGTACTATGTGACTCCAAACTGGAT 600

Qy 201 ILehTserHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 601 ATAACAAGCCACAAGAAAGTATATCAATCGTTGACGATGAAGAAAGAACGAGGACGC 660

Qy 221 HisHisLeuPheLeu 225

Db 661 CACCATCTGTTCTT 675

## RESULT 12

US-10-844-064A-1

Sequence 1, Application US/10844064A

Publication No. US20050149994A1

GENERAL INFORMATION:

APPLICANT: Bevis, Brooke

APPLICANT: Gluck, Benjamin

TITLE OF INVENTION: RAPIDLY MATURING FLUORESCENT PROTEINS AND METHODS FOR USING THE S

FILE REFERENCE: 092234-9006

CURRENT APPLICATION NUMBER: US/10/844,064A

CURRENT FILING DATE: 2004-05-11

PRIOR APPLICATION NUMBER: PCT/US02/40539

PRIOR FILING DATE: 2002-12-18

PRIOR APPLICATION NUMBER: US 60/341,723

PRIOR FILING DATE: 2001-12-19

NUMBER OF SEQ ID NOS: 3

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 678

TYPE: DNA

ORGANISM: Discosoma

US-10-844-064A-1

## Alignment Scores:

Pred. No.: 6,83e-143 Length: 678

Score: 1214.00 Matches: 225

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-844-064A-1 (1-678)

Qy 1 MetLArgSerSerIyBaenValIleIyGluPheMetAlaPheIyValaTgMetGluGly 20

Db 1 ATGAGGCTTCCTCAAGAAAGTATATCAAGAGATTCAATGAGGTTAAAGTTCCGATGAAAGCA 60

Qy 21 ThrValaAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 61 ACGGTCAATGGGACAGGTTTGAATAATGAAGGCGAAGAGGGGAGGCCATATGAAGGC 120

Qy 41 HisAsnThrValIyLeuLysValThrLysGlyGlyProLeuProPheAlaTfPaspIle 60

Db 121 CACATATCCGTAAAGCTTAAGTAAAGTAAACCAAGGGGAGACTTGTCCATTTGCTTGGGATATT 180

Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValIyHisPheProIaAspIlePro 80

Db 181 TTGTCACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACTCCGCGCATATCA 240

Qy 81 AepTyrIyLysLeuSerPheProGluGlyPheIyTfPgluaTgValMetAsnPheGlu 100

Db 241 GACTATATAAAAGCTGTCTTCCTGAAAGATTAAATGGGAAAGGCTCATCACTTTGAA 300

Qy 101 AaspGlyGlyValIValThrValThrGlnAspSerSerLeuGlnaAspGlyCyAspPheIleTyr 120

Db 301 GACGGTGGCGTGTACTGTAAACCCAGATTCCAGTTGCAAGATGGCTGTTCATCTAC 360

Qy 121 LysValIySphelIegIyValaAsnPheProSerAspGlyProValMetGlnLysLeuThr 140

Db 361 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCAGATGACCTGTTATGCAAAAGAGCA 420

Qy 141 MetGlyTfPgluaIAserThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 421 ATGGGCTGGGAAGCAGACACTGAGCGTTTGTATCTCTCGATGGCGTGTGAAAGGAGAG 480

Qy 161 ILehIySphalaleuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 481 ATTCAATAAAGCTGTGAAGCTAAAGACGGTGTCTATTAAGTTGAATTCAAAAGATTT 540

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValaAspSerLysLeuAsp 200

Db 541 TACATGGCAAGAAAGCCGTGTGACGTAACGAGGTACTATGTGACTCCAAACTGGAT 600

Qy 201 ILehTserHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 601 ATAACAAGCCACAAGAAAGTATATCAATCGTTGACGATGAAGAAAGAACGAGGACGC 660

Qy 221 HisHisLeuPheLeu 225

Db 661 CACCATCTGTTCTT 675

## RESULT 13

US-10-931-304-2

Sequence 2, Application US/10931304

Publication No. US20050196768A1

GENERAL INFORMATION:

APPLICANT: Tsien, Roger

APPLICANT: Campbell, Robert

TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT

FILE REFERENCE: 39754-0831CP2CP3

CURRENT APPLICATION NUMBER: US/10/931,304

CURRENT FILING DATE: 2004-08-30

PRIOR APPLICATION NUMBER: 10/209,208

PRIOR FILING DATE: 2002-07-29

PRIOR APPLICATION NUMBER: 10/121,258

PRIOR FILING DATE: 2002-04-10

PRIOR APPLICATION NUMBER: 09/866,538

PRIOR FILING DATE: 2001-05-24

PRIOR APPLICATION NUMBER: 09/794,308

PRIOR FILING DATE: 2001-02-26

NUMBER OF SEQ ID NOS: 110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 678

TYPE: DNA

ORGANISM: Discosoma sp.

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(678)

OTHER INFORMATION: wild-type Dared  
US-10-931-304-2

Alignment Scores:

Pred. No.:	6,83e-143	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-006-922a-12 (1-225) x US-10-931-304-2 (1-678)

QY	1	MeArGSeSerLyAsnValIleYsgLuphMeArGpHeLyValArgMeGluGly	20
DB	1	ATGAGGTTTCCAGAAATGTTATCAAGAGTTCAAGAGTTTAAGTTCCGATGGAGA	60
QY	21	ThrValaAngLYHsgLuphGluIleGluGlyGluGlyArgProTYrGluGly	40
DB	61	ACGGTCATGGGCAAGATTGAAATAGAGCGAAGAGGGGAGGCGCATACGAGGC	120
QY	41	HisAenThrValIleYsgLuphMeArGpHeLyValArgMeGluGly	60
DB	121	CACATACCGTAAGCTTAAGTAACCAAGGGGAGCTTGGCATTTGGCTGGAGATT	180
QY	61	LeuSerProGlnPheGlnTYrGlySerLYsValTYrValLYsHsProAlaAspIlePro	80
DB	181	TTGTCCACCAATTTCAAGTATGGAAGCAAGTATATGTCAGACCTCGCGACATACCA	240
QY	81	AspTYrLYsLYsLeuSerPheProGluGlyPheLYrTPGluArgValIleAsnPheGlu	100
DB	241	GACTATTAAGGCTGTCAATTTCTGAAGATTAAATGGAAAGGGTCATGAACTTTGAA	300
QY	101	AspGlyGlyValIleThrValIleThrGlnAspSerLeuGlnAspGlyCYsPheIleTYr	120
DB	301	GACGGTGGCGGTGTAAGTAAAGCAAGATTCAGATTCAGAGATGGCTGTTCACTAC	360
QY	121	LYsValIleYsgLuphMeArGpHeLyValArgMeGluGly	140
DB	361	AAAGTCAGATTCAATGGCGTGAACCTTCTTCGATGACCTGTTATGCAAAAGAGACA	420
QY	141	MeGlyTPGluAlaSerThrGluArgLeuTYrProArgAspGlyValIleLYsGlyGlu	160
DB	421	ATGGCTGGGAGACCAAGCACTGAGCGTTGTATCCCTGATGCGCTGTTGAAAGAGAG	480
QY	161	IleHsLYsValIleLYsLeuLYsAspGlyGlyHsLYrLeuValGluPheLYsSerIle	180
DB	481	ATTCAATTAAGGCTGTGAAGCTGAAGACGGTGTGCTATTAAGTTGAATTCAAAAGTATT	540
QY	181	TYrMeAlaLYsLYsProValGlnLeuProGlyTYrTYrValAspSerLYsLeuAsp	200
DB	541	TCAATGGCAAGAGCTGTGACGCTACCAAGGTTACTATGTTGACCTCAAACTGGAT	600
QY	201	IleHsSerHsAngLYsAspTYrThrIleValGluGlnTYrGluArgThrGluGlyArg	220
DB	601	ATTAACAAGCACAAGAAAGATTAACAATCGTTGAGCAAGATGAAGAAAGAGGAGCGC	660
QY	221	HisHsLYsLeuPheLeu	225
DB	661	CACCATCTGTCTCTT	675

RESULT 14

US-09-999-745-66  
Sequence 66, Application US/09999745  
Patent No. US20020157120A1  
GENERAL INFORMATION:  
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: Telen, Roger Y.  
APPLICANT: Baird, Geoffrey  
TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS  
FILE REFERENCE: REGEN1470-1  
CURRENT APPLICATION NUMBER: US/09/999,745  
CURRENT FILING DATE: 2001-10-23

PRIOR APPLICATION NUMBER: 09/316,920  
PRIOR FILING DATE: 1999-05-21  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 66  
LENGTH: 859  
TYPE: DNA  
ORGANISM: Drosophila sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54) .. (731)  
US-09-999-745-66

Alignment Scores:

Pred. No.:	9.49e-143	Length:	859
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-006-922a-12 (1-225) x US-09-999-745-66 (1-859)

QY	1	MeArGSeSerLyAsnValIleYsgLuphMeArGpHeLyValArgMeGluGly	20
DB	54	ATGAGGTTTCCAGAAATGTTATCAAGAGTTCAAGAGTTTAAGTTCCGATGGAGA	113
QY	21	ThrValaAngLYHsgLuphGluIleGluGlyGluGlyArgProTYrGluGly	40
DB	114	ACGGTCATGGGCAAGATTGAAATAGAGCGAAGAGGGGAGGCGCATACGAGGC	173
QY	41	HisAenThrValIleYsgLuphMeArGpHeLyValArgMeGluGly	60
DB	174	CACATACCGTAAGCTTAAGTAACCAAGGGGAGCTTGGCATTTGGCTGGAGATT	233
QY	61	LeuSerProGlnPheGlnTYrGlySerLYsValTYrValLYsHsProAlaAspIlePro	80
DB	234	TTGTCCACCAATTTCAAGTATGGAAGCAAGTATATGTCAGACCTCGCGACATACCA	293
QY	81	AspTYrLYsLYsLeuSerPheProGluGlyPheLYrTPGluArgValIleAsnPheGlu	100
DB	294	GACTATTAAGGCTGTCAATTTCTGAAGATTAAATGGAAAGGGTCATGAACTTTGAA	353
QY	101	AspGlyGlyValIleThrValIleThrGlnAspSerLeuGlnAspGlyCYsPheIleTYr	120
DB	354	GACGGTGGCGGTGTAAGTAAAGCAAGATTCAGATTCAGAGATGGCTGTTCACTAC	413
QY	121	LYsValIleYsgLuphMeArGpHeLyValArgMeGluGly	140
DB	414	AAAGTCAGATTCAATGGCGTGAACCTTCTTCGATGACCTGTTATGCAAAAGAGACA	473
QY	141	MeGlyTPGluAlaSerThrGluArgLeuTYrProArgAspGlyValIleLYsGlyGlu	160
DB	474	ATGGCTGGGAGACCAAGCACTGAGCGTTGTATCCCTGATGCGCTGTTGAAAGAGAG	533
QY	161	IleHsLYsValIleLYsLeuLYsAspGlyGlyHsLYrLeuValGluPheLYsSerIle	180
DB	534	ATTCAATTAAGGCTGTGAAGCTGAAGACGGTGTGCTATTAAGTTGAATTCAAAAGTATT	593
QY	181	TYrMeAlaLYsLYsProValGlnLeuProGlyTYrTYrValAspSerLYsLeuAsp	200
DB	594	TCAATGGCAAGAGCTGTGACGCTACCAAGGTTACTATGTTGACCTCAAACTGGAT	653
QY	201	IleHsSerHsAngLYsAspTYrThrIleValGluGlnTYrGluArgThrGluGlyArg	220
DB	654	ATTAACAAGCACAAGAAAGATTAACAATCGTTGAGCAAGATGAAGAAAGAGGAGCGC	713
QY	221	HisHsLYsLeuPheLeu	225
DB	714	CACCATCTGTCTCTT	728

RESULT 15

US-09-866-538-11

Sequence 11, Application US/09866538  
Publication NO. US20030032088A1  
GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSJEN, Roger  
APPLICANT: Campbell, Robert  
TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
FILE REFERENCE: REGEN1530-2  
CURRENT APPLICATION NUMBER: US/09/866,538  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 11  
LENGTH: 859  
TYPE: DNA  
ORGANISM: *Dicosoma sp.*  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54)..(731)  
US-09-866-538-11

Alignment Scores:  
Pred. No.: 9,49e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922A-12 (1-225) x US-09-866-538-11 (1-859)

QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20  
DB 54 ATGAGGCTCTTCAAGAAATGATCAAGAGATTCAAGAGTTAAAGTTCCGATGAAAGGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 114 ACGGTCAATGGGCGACGAGTTTGAATATGAAAGCGAAGAGAGGGAGGCCATACGAAGGC 173  
QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 174 CACAAATACCGTAAGCTTAAGTAACCAAGGGGGACCTTGGCCATTGGCTTGGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
DB 234 TTGTCCACCACAAATTCAGTATGAAAGCAAGGTATATGCAAGCACCTGCCGACATACCA 293  
QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTyrGluArgValMetAsnPheGlu 100  
DB 294 GACTATATAAAGCTGTCTATTCCTGAAGAGATTAAATGGAAAGGTCATGAACTTTGAA 353  
QY 101 AspGlyGlyValIleThrValThrGlnAspSerLeuGlnAspGlyCysPheIleTyr 120  
DB 354 GACCGTGGCGCTGTTACTGTAAACCAAGGATTCAGTTTGACAGATGGCTGTTTCATCTAC 413  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140  
DB 414 AAGGTCAAGTTCATGGCGGTGAACCTTCTCCAGTGAACCTGTTATGCAAAAGAAACA 473  
QY 141 MetGlyTyrGluIleAspThrGluArgLeuTyrProArgAspGlyValIleuIysGlyIle 160  
DB 474 ATGGGCTGGGAAGCCAGACACTGACGTTTGTATCTCTCGATGGCGTGTGAAGAGAGAG 533  
QY 161 IleHisIysValIleuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180  
DB 534 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGATTAACCTAAGTGAATCAAAAGTATT 593  
QY 181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200  
DB 594 TACATGGCAAAAGAGCCGTGACACTCAAGGTACTATATGTGACTCCAAACTGAT 653  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB 654 ATTAAGACCAACAGAGACTATACATCGTTGAGCAGTATGAAGAACCGAGGAGCGC 713

QY 221 HisHisLeuPheLeu 225

DB 714 CACCATCTGTCCCTT 728

Search completed: January 12, 2006, 12:40:06  
Job time : 797 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - nucleic search, using frame\_plus.p2n model

Run on: January 12, 2006, 10:16:32 | Search time 393 Seconds  
(without alignments)  
475.466 Million cell updates/sec

Title: US-10-006-922A-12

Perfect score: 1214  
Sequence: 1 NRSSKVIKEFMRFKVRMEG.....EDYIVQEYERTGRHHLFL 225

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus.model -DEV=rlp  
-Q=/cgn2\_1/USPTO.epool\_p/US10006922/rnatc\_10012006\_162410\_19008/app\_query.fasta\_1.391  
-DB=Published Applications NA.New -Qfmt=faecap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdt -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10006922 @CGN 1.1.184 @rnatc\_10012006\_162410\_19008  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPEXT=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA.New:

- 1: /cgn2\_6/prodata/1/pubpna/us08\_NEW\_PUB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/us06\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/us07\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/us09\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/us10\_NEW\_PUB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/us11\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/us11\_NEW\_PUB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/us11\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/us60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	100.0	678	6 US-10-209-208-2	Sequence 2, Appl1
2	1214	100.0	678	7 US-11-218-880-2	Sequence 2, Appl1
3	1210	99.7	681	6 US-10-209-208-3	Sequence 23, Appl1
4	1210	99.7	681	6 US-10-209-208-23	Sequence 2, Appl1
5	1210	99.7	681	7 US-11-218-880-3	Sequence 3, Appl1
6	1210	99.7	681	7 US-11-218-880-23	Sequence 23, Appl1
7	1186	97.7	6706	6 US-10-655-872-4	Sequence 4, Appl1
8	1186	97.7	7927	6 US-10-655-872-7	Sequence 7, Appl1

9	1160	95.6	678	6 US-10-209-208-5	Sequence 5, Appl1
10	1160	95.6	678	7 US-11-218-880-5	Sequence 5, Appl1
11	1121	92.3	681	6 US-10-209-208-7	Sequence 7, Appl1
12	1121	92.3	681	7 US-11-218-880-7	Sequence 7, Appl1
13	1021	84.1	678	6 US-10-209-208-9	Sequence 9, Appl1
14	1021	84.1	678	7 US-11-218-880-9	Sequence 9, Appl1
15	1012	83.4	678	6 US-10-209-208-80	Sequence 80, Appl1
16	486.5	40.1	1079	7 US-11-179-411-15	Sequence 15, Appl1
17	486.5	40.1	1079	7 US-11-179-766-15	Sequence 15, Appl1
18	474.5	39.1	1104	7 US-11-179-411-30	Sequence 30, Appl1
19	474.5	39.1	1104	7 US-11-175-766-30	Sequence 30, Appl1
20	474.5	39.1	1279	7 US-11-179-411-31	Sequence 31, Appl1
21	474.5	39.1	1279	7 US-11-175-766-31	Sequence 31, Appl1
22	468.5	38.6	6679	6 US-10-521-768-2	Sequence 1, Appl1
23	468.5	38.6	8251	6 US-10-521-768-3	Sequence 1, Appl1
24	468.5	38.6	10369	6 US-10-521-768-3	Sequence 2, Appl1
25	468.5	38.6	13535	6 US-10-521-768-3	Sequence 3, Appl1
26	422.5	34.8	860	7 US-11-179-411-24	Sequence 24, Appl1
27	422.5	34.8	860	7 US-11-175-766-24	Sequence 24, Appl1
28	422.5	34.8	864	7 US-11-179-411-23	Sequence 23, Appl1
29	422.5	34.8	864	7 US-11-179-411-26	Sequence 26, Appl1
30	422.5	34.8	864	7 US-11-175-766-23	Sequence 23, Appl1
31	422.5	34.8	864	7 US-11-175-766-26	Sequence 26, Appl1
32	422.5	34.8	873	7 US-11-179-411-25	Sequence 25, Appl1
33	422.5	34.8	873	7 US-11-175-766-25	Sequence 25, Appl1
34	219.5	18.1	10122	6 US-10-161-408-8	Sequence 8, Appl1
35	219.5	18.1	10549	6 US-10-161-408-3	Sequence 3, Appl1
36	219.5	18.1	12592	6 US-10-161-408-6	Sequence 6, Appl1
37	210.5	17.3	716	6 US-10-209-208-21	Sequence 21, Appl1
38	210.5	17.3	716	6 US-11-218-880-21	Sequence 21, Appl1
39	210.5	17.3	717	7 US-11-089-551A-12	Sequence 12, Appl1
40	210.5	17.3	720	7 US-11-089-551A-39	Sequence 39, Appl1
41	210.5	17.3	7239	7 US-11-089-551A-45	Sequence 45, Appl1
42	209.5	17.3	4773	7 US-11-214-613-32	Sequence 32, Appl1
43	209.5	17.3	5164	7 US-11-214-613-36	Sequence 36, Appl1
44	204.5	16.8	1539	7 US-11-032-236-5	Sequence 5, Appl1
45	204.5	16.8	3660	7 US-11-032-236-3	Sequence 3, Appl1

#### ALIGNMENTS

RESULT 1  
US-10-209-208-2  
; Sequence 2, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey, Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: *Drosophila* sp.  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(678)  
; OTHER INFORMATION: wild-type DsRed  
US-10-209-208-2  
Alignment Scores:

Pred. No.: 2,23e-133 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-209-208-2 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
 Db 1 ATGAGGCTCTCCAAAGATGTATCAAGAGATTCAAGAGTTAAAGTTCCGATCGATGAAGGA 60  
 QY 21 ThrValAsnGlyHsGluPheGluIleGluGlyGluGlyValArgProTyrgLugly 40  
 Db 61 ACGGTCAATGGGACCGAGTTTGAATATGAAAGCGGAGAGGGGAGGCCATACGAAGGC 120  
 QY 41 HisAsnThrValLySLeuLySValThrLySGlyProLeuProPheAlaTrpAspIle 60  
 Db 121 CACAAATACCGTAAAGCTTAAGSTAACAAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
 QY 61 LeuSerProGlnPheGlnTyrgLySerLySValTyrgValLySHisProAlaAspIlePro 80  
 Db 181 TTGTCCACCACAAATTCAGTATGGAAGCAAGGTATATGCAAGCACTTCGCCACATGCCA 240  
 QY 81 AspTyrlLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
 Db 241 GACTATMAAAGCTGTGATCTTCCTGAAAGATTAAATGGAAAGGGTCATGAACTTTGAA 300  
 QY 101 AspGlyGlyValIleThrValThrGlnAspSerSerLeuGlnAspGlyCySPhelIeTy 120  
 Db 301 GACGGTGGCGTGTACTGTAACCAAGATTCGAGTTTGACAGATGGCTGTTCATCTAC 360  
 QY 121 LySValLySPhelIeGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
 Db 361 AAGGTCAAGTTCATTTGGCGTGAACCTTCCATCCAGTGAACCTGTATGCAAAAGACCA 420  
 QY 141 MetGlyTrpGluIleAserThrGluArgLeuTyrgProArgAspGlyValIleLySLeu 160  
 Db 421 ATGGGCTGGGAAGCCGACCTGAGCGCTTTGTATCTCGTAAAGCGCTGTAAAGGAGAG 480  
 QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHsTyrgLeuValGluPheLySLeu 180  
 Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGCTATTACCTAGTTGAATCAAAAGTATT 540  
 QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrgTyrgValAspSerLySLeuAsp 200  
 Db 541 TACATGGCAAAAGAGCCGTGACCTACCAAGGTACTACTATGTGACTCCAAACTGGAT 600  
 QY 201 IleThrSerHisAsnGluAspTyrgThrIleValGluGlnTyrgIuArgThrGluGlyArg 220  
 Db 601 ATAAACAAGCCACAACGAAGCTATATCAATCGTTAGCAGTATGAAGAACCGAGGAGCGC 660  
 QY 221 HisHisLeuPheLeu 225  
 Db 661 CACCATCTGTCTT 675  
 RESULT 2  
 US-11-218-880-2  
 ; Sequence 2, Application US/11218880  
 ; Publication No. US20060003420A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Telen, Roger  
 ; APPLICANT: Campbell, Robert  
 ; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
 ; FILE REFERENCE: UC083.1CP2CPI  
 ; CURRENT APPLICATION NUMBER: US/11/218.880  
 ; PRIOR FILING DATE: 2005-09-01  
 ; PRIOR APPLICATION NUMBER: US/10/121.258  
 ; PRIOR FILING DATE: 2002-04-10  
 ; PRIOR APPLICATION NUMBER: 09/794,308  
 ; PRIOR FILING DATE: 2001-02-26

; PRIOR APPLICATION NUMBER: 09/866,538  
 ; PRIOR FILING DATE: 2001-05-24  
 ; NUMBER OF SEQ. ID NOS: 78  
 ; SOFTWARE: FastSeq For Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 678  
 ; TYPE: DNA  
 ; ORGANISM: Diatomsoma sp.  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(678)  
 ; OTHER INFORMATION: wild-type Dared  
 US-11-218-880-2

#### Alignment Scores:

Pred. No.: 2,23e-133 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-006-922a-12 (1-225) x US-11-218-880-2 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
 Db 1 ATGAGGCTCTCCAAAGATGTATCAAGAGATTCAAGAGTTAAAGTTCCGATCGATGAAGGA 60  
 QY 21 ThrValAsnGlyHsGluPheGluIleGluGlyGluGlyValArgProTyrgLugly 40  
 Db 61 ACGGTCAATGGGACCGAGTTTGAATATGAAAGCGGAGAGGGGAGGCCATACGAAGGC 120  
 QY 41 HisAsnThrValLySLeuLySValThrLySGlyProLeuProPheAlaTrpAspIle 60  
 Db 121 CACAAATACCGTAAAGCTTAAGSTAACAAAGGGGGGACCTTTGCCATTTGCTTGGGATATT 180  
 QY 61 LeuSerProGlnPheGlnTyrgLySerLySValTyrgValLySHisProAlaAspIlePro 80  
 Db 181 TTGTCCACCACAAATTCAGTATGGAAGCAAGGTATATGTCAGACACCTTCGCCACATGCCA 240  
 QY 81 AspTyrlLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
 Db 241 GACTATMAAAGCTGTGATCTTCCTGAAAGATTAAATGGAAAGGGTCATGAACTTTGAA 300  
 QY 101 AspGlyGlyValIleThrValThrGlnAspSerSerLeuGlnAspGlyCySPhelIeTy 120  
 Db 301 GACGGTGGCGTGTACTGTAACCAAGATTCGAGTTTGACAGATGGCTGTTCATCTAC 360  
 QY 121 LySValLySPhelIeGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
 Db 361 AAGGTCAAGTTCATTGGCGGAACTTTCCCTCGATGAGACTGTATGCAAAAGACCA 420  
 QY 141 MetGlyTrpGluIleAserThrGluArgLeuTyrgProArgAspGlyValIleLySLeu 160  
 Db 421 ATGGGCTGGGAAGCCGACCTGAGCGCTTTGTATCTCGTATGAGCGGTGTAAAGGAGAG 480  
 QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHsTyrgLeuValGluPheLySLeu 180  
 Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGCTATTACCTAGTTGAATCAAAAGTATT 540  
 QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrgTyrgValAspSerLySLeuAsp 200  
 Db 541 TACATGGCAAAAGAGCTGTGACCTACCAAGGTACTACTATGTGACTCCAAACTGGAT 600  
 QY 201 IleThrSerHisAsnGluAspTyrgThrIleValGluGlnTyrgIuArgThrGluGlyArg 220  
 Db 601 ATAAACAAGCCACAACGAAGCTATATCAATCGTTGACAGATATGAAGAACCGAGGAGCGC 660  
 QY 221 HisHisLeuPheLeu 225  
 Db 661 CACCATCTGTCTT 675  
 RESULT 3



```
Qy 101 AspglyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
    |||
Db 304 GACGGCGCGGTGATCCGTGAACCAAGACCTCTCTCCAGAGAGCGCTTCACTTAC 363
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLeuThr 140
    |||
Db 364 AAGGTGAAGTTCAATCGCGGTGAATCTCCCTCCGACGGCCCCGTAAATGACAGAGACC 423
Qy 141 MetGlyTTPGluLAserThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
    |||
Db 424 ATGGCTGGAGGCGCTCCACCGAGGCGCTGTACCCCGACGGCGGTCTGAAGGGCAG 483
Qy 161 LLeHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
    |||
Db 484 ATCCACAGAGCGCTGACACTGAAGACCGCGGCACTACTCTGTGTGAGTTCAAGTCCATC 543
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
    |||
Db 544 TACATGGCCAGAGAGCCCGTGCAGCTGCCGCTACTACTAGTGAATCCAGCTGAC 603
Qy 201 LLeHisSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
    |||
Db 604 ATCACTCCCAACAAGAGACTACACATCGTGAAGACGTACAGCGCACCGAGGGCGCGC 663
Qy 221 HisHisLeuPheLeu 225
    |||
Db 664 CACCACTGTTCCTG 678

RESULT 5
US-11-218-880-3
; Sequence 3, Application US/11218880
; Publication No. US20060003420A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CPI
; CURRENT APPLICATION NUMBER: US/11/218,880
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US/10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian
; OTHER INFORMATION: codon usage
US-11-218-880-3

Alignment Scores:
Pred. No.: 6,67e-133 Length: 681
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-3 (1-681)
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetAlaArgPheLysValArgMetGluGly 20
    ::|
Db 4 GTGGCTTCCTCCAAAGAGATCAAGAGATTCAATGCGCTTCAAGAGTCCGATGAGAGGC 63
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
```

```
Db 64 ACCGTGAACGGCCACGAGTTGAGATGAGGGGAGGGCCGCGCCCTTACAGAGGC 123
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheIleATTPAspIle 60
    |||
Db 124 CACAACACCGCTGAAGCTGAAGGTGAACCAAGGGCGCCCTTCCCTTCCGCTTGGAGATC 183
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
    |||
Db 184 CTGCCCCCGAGTTCCAGTACCGCTCCAGAGGTGACGTGAAGCACCCCGCGACATCCCC 243
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTTPGluArgValMetAsnPheGlu 100
    |||
Db 244 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCGAG 303
Qy 101 AspglyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
    |||
Db 304 GACGGCGCGGTGATCCGTGAACCAAGACCTCTCTCCGACGGCCCCGTAAATGACAGAGACC 363
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLeuThr 140
    |||
Db 364 AAGGTGAAGTTCAATCGCGGTGAATCTCCCTCCGACGGCCCCGTAAATGACAGAGACC 423
Qy 141 MetGlyTTPGluLAserThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
    |||
Db 424 ATGGCTGGAGGCGCTCCACCGAGGCGCTGTACCCCGACGGCGGTCTGAAGGGCAG 483
Qy 161 LLeHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
    |||
Db 484 ATCCACAGAGCGCTGACACTGAAGACCGCGGCACTACTCTGTGTGAGTTCAAGTCCATC 543
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
    |||
Db 544 TACATGGCCAGAGAGCCCGTGCAGCTGCCGCTACTACTAGTGAATCCAGCTGAC 603
Qy 201 LLeHisSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
    |||
Db 604 ATCACTCCCAACAAGAGACTACACATCGTGAAGACGTACAGCGCACCGAGGGCGCGC 663
Qy 221 HisHisLeuPheLeu 225
    |||
Db 664 CACCACTGTTCCTG 678

RESULT 6
US-11-218-880-23
; Sequence 23, Application US/11218880
; Publication No. US20060003420A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CPI
; CURRENT APPLICATION NUMBER: US/11/218,880
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US/10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian
; OTHER INFORMATION: codon usage
US-11-218-880-23

Alignment Scores:
```



Pred. No.: 6,678-133 Length: 681  
 Score: 1210.00 Matches: 224  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.56% Mismatches: 0  
 Query Match: 99.67% Indels: 0  
 DB: 7 Gaps: 0

US-10-006-922a-12 (1-225) x US-11-218-880-23 (1-681).

```

QY 1 MetAAGSerSerLyAaenValIleLygLuPhemeCArPhelyVaIArgMeGlUgLy 20
DB 4 GTGCCCTCTCCAGAAAGCTCATCAAGATTCAAGCTTCAAGTCCGATGAGAGG 63
QY 21 ThrValaengLyHsAgLuPhelunIleGlUgLyglUgLyglUgLyArgProTyrgLgLy 40
DB 64 ACCGGAACGGCCAGAGTTGAGATCGAGGGCGAGGGCGGCCGCCCTTACAGAGG 123
QY 41 HisAenThrValLyLeuLyVaIThrLygLyglYProLeuProPhaIaITraAapIle 60
DB 124 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCCTTGGCGAGACATC 183
QY 61 LeuSerProGlnPhelunTyrgLySerLyVaIlyrValLyHsAProIlaAapIlePro 80
DB 184 CTGTCCCCCAATTCAGTACAGGCTTCAAGGTGTACGTGAACACCCCGCCAGCATCCC 243
QY 81 AapTyLyLyLeuSerPheProGluGlyPheLySTrPGLuArgValMetAanPheGlu 100
DB 244 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTATGAATTCGAG 303
QY 101 AapGlyLyValValThrValThrGlnAapSerSerLeuGlnAapGlyCyAaPheIleTy 120
DB 304 GACGGCGGGGTGTGACCGGTGACCCAGGACTCTCTCCGTGACAGAGCGGCTTCAATAC 363
QY 121 LySValLyPheIleGlyValAanPheProSerAapGlyProValMeGlunLySlyThr 140
DB 364 AAGGTGAATCATCGGCGTGAATCTTCCCTCGAGCGCCCGGTATGAGAAAGAAC 423
QY 141 MetGlyTrpGluAaSerThrGluArgLeuTyProArgAapGlyValLeuLygLyGlu 160
DB 424 ATGGGTGGAGAGGCTCCACCGAGGCGCTGTACCCCGGAGCGGCGTGAAGGGCGAG 483
QY 161 IleHsLyAAlaLeuLyLeuLyAaPGLyGlyHsITyLeuValGluPhelySerIle 180
DB 484 ATCCACAAAGCCCTGAAGCTGAAGAGCGGGCGCACTACCTGTGTGAAGTTCATAC 543
QY 181 TyrMetAlaLyLyAProValGlnLeuProGlyTyTyTyTyValAapSerLySleuAap 200
DB 544 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGTACTACTACGTGAGCTTCAAGCTGAG 603
QY 201 IleThrSerHisAenGluAapTyThrIleValGluGlnTyrgLgLyArgThrgLgLyArg 220
DB 604 ATCACTCCCAACAAGAGCTACATCATGTGTGAGCAGTACGAGGCAACGAGGGCGCG 663
QY 221 HisHsLeuPheLeu 225
DB 664 CACCACTGTCTCTG 678

```

RESULT 7  
 US-10-655-872-4  
 ; Sequence 4, Application US/10655872  
 ; Publication No. US20050251872A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bear, et al.  
 ; TITLE OF INVENTION: Lentiviral Vectors, Related Reagents, and Methods of Use Thereof  
 ; FILE REFERENCE: 0492611-0512  
 ; CURRENT FILING DATE: 2003-09-05  
 ; SOFTWARE: Patent in version 3.2  
 ; SEQ ID NO 4  
 ; LENGTH: 6706  
 ; TYPE: DNA  
 ; ORGANISM: Artificial

FEATURE:  
 ; OTHER INFORMATION: Lentiviral vector sequence  
 ; US-10-655-872-4

Alignment Scores:

Pred. No.: 1e-128 Length: 6706  
 Score: 1186.00 Matches: 219  
 Percent Similarity: 98.22% Conservative: 2  
 Best Local Similarity: 97.33% Mismatches: 4  
 Query Match: 97.69% Indels: 0  
 DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-655-872-4 (1-6706)

```

QY 1 MetAAGSerSerLyAaenValIleLygLuPhemeCArPhelyVaIArgMeGlUgLy 20
DB 2778 ATGGCTCTCCCGAAGAGTCAACCAAGTTCATCAAGCTTCAAGTCCGATGAGAGG 2837
QY 21 ThrValaengLyHsAgLuPhelunIleGlUgLyglUgLyglUgLyArgProTyrgLgLy 40
DB 2838 ACCGGAACGGCCAGAGTTGAGATCGAGGGCGAGGGCGGCCGCCCTTACAGAGG 2897
QY 41 HisAenThrValLyLeuLyVaIThrLygLyglYProLeuProPhaIaITraAapIle 60
DB 2898 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCCTTGGCGAGACATC 2957
QY 61 LeuSerProGlnPhelunTyrgLySerLyVaIlyrValLyHsAProIlaAapIlePro 80
DB 2958 CTGTCCCCCAATTCAGTACAGGCTTCAAGGTGTACGTGAAGACCCCGCCAGCATCCC 3017
QY 81 AapTyLyLyLeuSerPheProGluGlyPheLySTrPGLuArgValMetAanPheGlu 100
DB 3018 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTATGAATTCGAG 3077
QY 101 AapGlyLyValValThrValThrGlnAapSerSerLeuGlnAapGlyCyAaPheIleTy 120
DB 3078 GACGGCGGGGTGTGACCGGTGACCCAGGACTCTCTCCGTGACAGAGCGGCTTCAATAC 3137
QY 121 LySValLyPheIleGlyValAanPheProSerAapGlyProValMeGlunLySlyThr 140
DB 3138 AAGGTGAATCATCGGCGTGAATCTTCCCTCGAGCGCCCGGTATGAGAAAGAAC 3197
QY 141 MetGlyTrpGluAaSerThrGluArgLeuTyProArgAapGlyValLeuLygLyGlu 160
DB 3198 ATGGGTGGAGAGGCTCCACCGAGGCGCTGTACCCCGGAGCGGCGTGAAGGGCGAG 3257
QY 161 IleHsLyAAlaLeuLyLeuLyAaPGLyGlyHsITyLeuValGluPhelySerIle 180
DB 3258 ACCCAAGAGCCCTGAAGCTGAAGAGCGGGCGCACTACCTGTGTGAAGTTCATAC 3317
QY 181 TyrMetAlaLyLyAProValGlnLeuProGlyTyTyTyTyValAapSerLySleuAap 200
DB 3318 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGTACTACTACGTGAGACCCCAAGCTGAG 3377
QY 201 IleThrSerHisAenGluAapTyThrIleValGluGlnTyrgLgLyArgThrgLgLyArg 220
DB 3378 ATCACTCCCAACAAGAGCTACATCATGTGTGAGCAGTACGAGGCAACGAGGGCGCG 3437
QY 221 HisHsLeuPheLeu 225
DB 3438 CACCACTGTCTCTG 3452

```

RESULT 8  
 US-10-655-872-7  
 ; Sequence 7, Application US/10655872  
 ; Publication No. US20050251872A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bear, et al.  
 ; TITLE OF INVENTION: Lentiviral Vectors, Related Reagents, and Methods of Use Thereof  
 ; FILE REFERENCE: 0492611-0512  
 ; CURRENT FILING DATE: 2003-09-05  
 ; NUMBER OF SEQ ID NOS: 39

	APPLICANT:	Campbell, Robert
	APPLICANT:	Geoffrey Baird
	TITLE OR INVENTION:	FLUORESCENT PROTEIN VARIANTS AND METHODS
	TITLE OF INVENTION:	FOR MAKING SAME
	FILE REFERENCE:	UC083.1CP2CP2
	CURRENT APPLICATION NUMBER:	US/10/209,208
	PRIOR FILING DATE:	2002-07-29
	PRIOR APPLICATION NUMBER:	10/121,258
	PRIOR FILING DATE:	2002-04-10
	PRIOR APPLICATION NUMBER:	09/866,538
	PRIOR FILING DATE:	2001-05-24
	PRIOR APPLICATION NUMBER:	09/794,308
	PRIOR FILING DATE:	2001-02-26
	NUMBER OF SEQ ID NOS:	80
	SOFTWARE:	FastSeq for Windows Version 4.0
	SEQ ID NO 5	
	LENGTH:	678
	TYPE:	DNA
	ORGANISM:	Artificial Sequence
	FEATURES:	
	OTHER INFORMATION:	Polynucleotide encoding Dared polypeptide variant "T1"
	OTHER INFORMATION:	"T1"
	US-10-209-208-5	
	Alignment Scores:	
	Score:	5.3e-127 Length: 678
	Percent Similarity:	1160.00 Matches: 216
	Best Local Similarity:	97.33% Conservative: 3
	Query Match:	96.00% Mismatches: 6
	DB:	95.55% Indels: 0
		Gaps: 0
	US-10-006-922A-12 (1-225) x US-10-209-208-5 (1-678)	
QY	1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly	20
DB	1 ATGGCTCTCTCCGAGGACGTCTATCAAGAGTTTCATCGCTTCAAGGTGCCTGCGCATGAGGCC	60
QY	21 ThrValAsnGIYHISgluPheGluIleGluGlyGluGluGlyArgProTYrgluGly	40
DB	61 TCGGTGAACGGCCACGAGATTGAGATCGAGGCGAGGGCGAGGGCCGCCCTCAGAGGGC	120
QY	41 HISasnthValIlysLeuLysValThrtYrsGlyGlyProLeuProPheAlaTPRpplle	60
DB	121 ACCGACGCCGCCAACTGAAGGTGACCAGGGCGGGCCCCCTTCGCTGGGACATC	180
QY	61 LeuSerProGlnPheGluITyrgIySerIysValTYrValIysSHisProAlaAppllePro	80
DB	181 CTGTCCCCCAAGTTCAGTAGCGGCTCCAGAGTGTACTGAAGCACCCCGCCGACATCCC	240
QY	81 AspTYrLyAllysLeuSerPheProGluGlyPheIysTPTGLuarGValMetAsnPheGlu	100
DB	241 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAATGGAGGCGCGTGAATACTTCGAG	300
QY	101 AsPGlyGlyValIvalThryalThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyrr	120
DB	301 GACGGGGGGGTGTGACCGTGACCAAGACATCTCCCTCGACGAGCGGCTCTCATCTAC	360
QY	121 LysValIysPheIIleGIVaIAsnpPheProSerAspGlyProValIMetGlnLYsLYThr	140
DB	361 AAGGTAAAGTATCGGGGTGAATCTTCCCTTCGACGAGCCCCCGTAAATGCAAGAGAACT	420
QY	141 MetGlyTYrGluuAlaserThrcIuarIgsueuTYrProArghAspGlyValLeuIysGlyLu	160
DB	421 ATGGGTGGGAAGCTCCACGAGGGCTGTAAACCCCGCGACGCGGTGAAGGGCGAG	480
QY	161 IIEHisLYsAlIleuLYsLeuLYsAspGlyGlyHIstYrIeuValGIupheIysSerIle	180
DB	481 ATCCAAGAACCCCTAAGCTGAAGGACGGCGCACTAAGTGGTGAATTCATCATC	540
QY	181 TYrMeAlALysLYsProValGInLeuProGlyTYrTYrTYrValAAPserLYsLeuAsp	200
DB	541 TACATGGCCAAAGACCCGTGACGTGCCGGGTCTACTACTACGTGACCTCAAGCTGGAC	600

```
QY 201 ILeHrSeRHiSaNgLuAaPpTyThRiLeVaIgLuGInTyRGUaRgThRgLuGlyArg 220
DB 601 ATCACTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCGCCGAGGCGCGC 660
QY 221 HSHSLeUpHeLeu 225
DB 661 CACCACTGTTCCTG 675

RESULT 10
US-11-218-880-5
/ Sequence 5, Application US/11218880
/ Publication No. US20060003420A1
/ GENERAL INFORMATION:
/ APPLICANT: Teien, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: UC083.1CP2CP1
/ CURRENT APPLICATION NUMBER: US/11/218,880
/ CURRENT FILING DATE: 2005-09-01
/ PRIOR APPLICATION NUMBER: US/10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Dated polypeptide variant
/ OTHER INFORMATION: "T1"
US-11-218-880-5

Alignment Scores:
Pred. No.: 5,3e-127 Length: 678
Score: 1160.00 Matches: 216
Percent Similarity: 97.33% Conservative: 3
Best Local Similarity: 96.00% Mismatches: 6
Query Match: 95.55% Indels: 0
DB: 7 Gaps: 0

US-10-006-922a-12 (1-225) x US-11-218-880-5 (1-678)
QY 1 MetArgSerSerLySaAnValIleLySGluPHeMetArgPHeLySaValArgMetGluGly 20
DB 1 ATGGCCTCCCGAAGAGCGTCATCAAGAGTTCATGCGCTTCAAGGCGCATGAGAGGC 60
QY 21 ThRValaSaNgLyHISgluPHeGluIleGluGlyGluGlyArgProTyRgLuGly 40
DB 61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCCGCGCCCTTACGAGGC 120
QY 41 HLeaAnThRValLySaLeuLySaValThRlyGlyGlyProLeuProPheLaATpAaPle 60
DB 121 ACCCAAGCCCGCAAGCTGAGAGGTGACCAAGGGCGGCCCTTCCCTTCCGCTGGACATC 180
QY 61 LeuSerProGlnPHeGlnTyRgLySerLySaValTyRValLySHsAProLaAaPlePro 80
DB 181 CTGTCCCGCCAGTTCAGTACGGCTCCAGGTGATGATGAAGACCCCGCCGATCCCC 240
QY 81 AspTyRlyLySaLeuSerPHeProGluGlyPHeLySTpGluArgValMetLaPHeGlu 100
DB 241 GACTTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGGTGAGAGCGCGTGAATGAATCTTGAG 300
QY 101 AspGlyGlyValValThRValThRglnAspSerSerLeuGlnAspGlyCySPheileTyR 120
DB 301 GACGGCGGCGTGTGACCGTGAACCAAGACTCTCTCCCTGACAGACGGCTCTTCATATAC 360
QY 121 LySaValLySaPHeileGlyValaSaPHeProSerAspGlyProValMetGlnLySaThR 140
```

```
DB 361 AAGGTGAAGTTCACTCGGCTGAACCTCCCTCCAGCGGCCCGGTAAATGCAAGAAGACT 420
QY 141 MetGlyTPRgLuJaSaSerThRgLuArgLeuTyRProArgAspGlyValLeuLySGly 160
DB 421 ATGGCTGTGAGAGGCTTCCACCGAGGCTGTACCCCGCCGAGCGGTGCTGTAAGAGGCGAG 480
QY 161 ILeHSLySaLaLeuLySaLeuLySaPpLyGlyHISTyRLeuValGluPHeLySerIle 180
DB 481 ATCCCAAGAGCGCCGTGAAGACTGAAGAGCGCGCACCTGCTGTGAGTTCAGATTCATC 540
QY 181 TyrMetAlaLySaProValGlnLeuProGlyTyRlyTyRValaSaPserLySaLeuAsp 200
DB 541 TACATGCCCAAGAACCCCTGTCACTGCTCCCGCTACTACTAGTGAATCCCAAGCTGAGAC 600
QY 201 ILeHrSeRHiSaNgLuAaPpTyThRiLeVaIgLuGInTyRGUaRgThRgLuGlyArg 220
DB 601 ATCACTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCGCCGAGGCGCGC 660
QY 221 HSHSLeUpHeLeu 225
DB 661 CACCACTGTTCCTG 675

RESULT 11
US-10-209-208-7
/ Sequence 7, Application US/10209208
/ Publication No. US2005024921A1
/ GENERAL INFORMATION:
/ APPLICANT: Teien, Roger
/ APPLICANT: Campbell, Robert
/ APPLICANT: Geoffrey Baird
/ TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
/ TITLE OF INVENTION: FOR MAKING SAME
/ FILE REFERENCE: UC083.1CP2CP2
/ CURRENT APPLICATION NUMBER: US/10/209,208
/ CURRENT FILING DATE: 2002-07-29
/ PRIOR APPLICATION NUMBER: 10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 80
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7
/ LENGTH: 681
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Dated polypeptide variant
/ OTHER INFORMATION: "dimer2"
US-10-209-208-7

Alignment Scores:
Pred. No.: 2,15e-122 Length: 681
Score: 1121.00 Matches: 207
Percent Similarity: 95.96% Conservative: 7
Best Local Similarity: 92.83% Mismatches: 9
Query Match: 92.34% Indels: 0
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-209-208-7 (1-681)
QY 3 SerSerLySaAnValIleLySGluPHeMetArgPHeLySaValArgMetGluGlyThRVal 22
DB 10 TCCTCCGAGAGAGCTCATCAAGAGTTCATGCGCTTCAAGGTGCGATGAGAGGCTCCGTG 69
QY 23 AsnGlyHISgluPHeGluIleGluGlyGluGlyArgProTyRgLuGlyHISaAn 42
DB 70 AACGGCCACGAGTTCAGATTCGAGGGCGAGGGCGAGGGCGGCCCTTACGAGGACCCAG 129
QY 43 ThRValLySaLeuLySaValThRlySGlyGlyProLeuProPheLaATpAaPleLeuSer 62
```

Db 130 ACCGCCAAGCTGAAGGTGACCAAGGCGGCCCCCTTCGCTTGAGACATCTGTC 189  
Qy 63 ProGlnPheGlnTYrGlySerIyValIyTyValIyShiSPRoAlaAspIleProAspTYr 82  
Db 190 CCCAGTTCGAGTACGGCTCCAGAGCGTACGTAGAACACCCCGCGACATCCCGACATAC 249  
Qy 83 LysIleuSerPheProGluGlyPheIySTrpgIuAqValIleAsnPheGluAspGly 102  
Db 250 AAGAACTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGACCTTGAGAGCGGC 309  
Qy 103 GLyValValIThrValIThrGlnAspSerSerIeuGlnAspGlyCysPheIleTYrIyVal 122  
Db 310 GCGCGGTGACCGTGAACCAAGACTCTCCCTGACGAGCGACGCTGATCTACAAAGGTG 369  
Qy 123 LysPheIleGlyValIAsnPheProSerAspGlyProValIleMetGlnIyLysIleThMetGly 142  
Db 370 AAGTTCGCGGACCAACTTCCCGCGAGCGCCCGGTAAAGCAGAAACACCATGGGC 429  
Qy 143 TrpGluIaSerThrGluAqIleuTYrProArgAspGlyValIleuLysGlyIuIleHis 162  
Db 430 TGGAGGCTTCACCGAGCGCTGTACCCCGCGACGCGTGTGAGAGCGAGATCCAC 489  
Qy 163 LysAlaIeuLysIeuLysAspGlyGlyIySTyIleuValIguPheLysSerIleTYrMet 182  
Db 490 CAGGCGCTGAAGCTGAAGAGCGGCGGCACTACCTGGTGAAGTTCAGACCATCTACATG 549  
Qy 183 AlaIyLysProValGlnIeuProGlyTYrTYrTYrValIAspSerIyLysIleuAspIleThr 202  
Db 550 GCCAAGAGCCCGTGCACCTCCCGCTACTACTACGTGACACCAACAGCTGACATCC 609  
Qy 203 SerIleAsnGluAspTYrThrIleValIguIuTYrGluAqIleuTYrGluIyArgHis 222  
Db 610 TCCCAACAAGAGACTACCATCTGTGAAACAATGACGCTCCGAGGCGCCGACAC 669  
Qy 223 LeuPheIeu 225  
Db 670 CTGTTCTCTG 678

RESULT 12  
US-11-218-880-7  
; Sequence 7, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; PRIOR FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding Dered polypeptide variant  
; OTHER INFORMATION: "dimer2"  
US-11-218-880-7

Alignment Scores:  
Pred. No.: 2.15e-122 Length: 681  
Score: 1121.00 Matches: 207  
Percent Similarity: 95.96% Conservatve: 7  
Best Local Similarity: 92.83% Mismatches: 9  
Query Match: 92.34% Indels: 0

DB: 7 Gaps: 0  
US-10-006-922a-12 (1-225) x US-11-218-880-7 (1-681)  
Qy 3 SerSerLysAsnValIleIyLysGluPheMetArgPheLysValIArgMetGluGlyIThrVal 22  
Db 10 TCTCCAGAGACGTATCAAAAGATTATGATGCTTCAAGGTGGCGCATGAGGCGTCCGTG 69  
Qy 23 AsnGlyIySTrpgIuIleGluGlyIguIyGluIyArgProTYrGluIyIyIleAsn 42  
Db 70 AACGGCACAAGTTCGAGATCGAAGGCGAGGCGCAAGGCGCCCTACGAGGCGACCCAG 129  
Qy 43 ThrValIyLysIeuLysValIThrIyGlyIyProIeuProPheIleATrPAAspIleLysSer 62  
Db 130 ACCGCCAAGCTGAAGGTGACCAAGGCGGCCCCCTTCGCTTGAGACATCTGTC 189  
Qy 63 ProGlnPheGlnTYrGlySerIyValIyTyValIyShiSPRoAlaAspIleProAspTYr 82  
Db 190 CCCAGTTCGAGTACGGCTCCAGAGCGTACGTAGAACACCCCGCGACATCCCGACATAC 249  
Qy 83 LysIleuSerPheProGluGlyPheIySTrpgIuAqValIleAsnPheGluAspGly 102  
Db 250 AAGAACTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGACCTTGAGAGCGGC 309  
Qy 103 GLyValValIThrValIThrGlnAspSerSerIeuGlnAspGlyCysPheIleTYrIyVal 122  
Db 310 GCGCGGTGACCGTGAACCAAGACTCTCCCTGACGAGCGACGCTGATCTACAAAGGTG 369  
Qy 123 LysPheIleGlyValIAsnPheProSerAspGlyProValIleMetGlnIyLysIleThMetGly 142  
Db 370 AAGTTCGCGGACCAACTTCCCGCGAGCGCCCGGTAAAGCAGAAACACCATGGGC 429  
Qy 143 TrpGluIaSerThrGluAqIleuTYrProArgAspGlyValIleuLysGlyIuIleHis 162  
Db 430 TGGAGGCTTCACCGAGCGCTGTACCCCGCGACGCGTGTGAGAGCGAGATCCAC 489  
Qy 163 LysAlaIeuLysIeuLysAspGlyGlyIySTyIleuValIguPheLysSerIleTYrMet 182  
Db 490 CAGGCGCTGAAGCTGAAGAGCGGCGGCACTACCTGATGAGATTCAGACCATCTACATG 549  
Qy 183 AlaIyLysProValGlnIeuProGlyTYrTYrTYrValIAspSerIyLysIleuAspIleThr 202  
Db 550 GCCAAGAGCCCGTGCACCTCCCGCTACTACTACGTGACACCAACAGCTGACATCC 609  
Qy 203 SerIleAsnGluAspTYrThrIleValIguIuTYrGluAqIleuTYrGluIyArgHis 222  
Db 610 TCCCAACAAGAGACTACCATCTGTGAAACAATGACGCTCCGAGGCGCCGACAC 669  
Qy 223 LeuPheIeu 225  
Db 670 CTGTTCTCTG 678

RESULT 13  
US-10-209-208-9  
; Sequence 9, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0

```

/ SEQ ID NO 5
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Dared polypeptide variant
/ OTHER INFORMATION: "MBFP1"
US-10-209-208-9

Alignment Scores:
Pred. No.: 1,366-110 Length: 678
Score: 1021.00 Matches: 192
Percent Similarity: 89.14% Conserved: 5
Beac Local Similarity: 86.88% Mismatches: 24
Query Match: 84.10% Indels: 0
DB: Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-9 (1-678)

QY 1 MetAgsSerSeLyAaenVal11LeYsGluPheMetAaPhelyValaLargMetGluY 20
Db 1 ATGGCTCTCTCCGAGAGCGCTATCAAGAGGTTCAAGCCCTTCAAGAGGTCGCAAGAGGC 60
QY 21 ThrValAaNG1VH1sGluPheGlu1LeGluGluY1GluY1GluY1AaProYrGluY 40
Db 61 TCGGTAAAGCCCAAGATTCCAGATTCAGAGGCGAAGGCGCGGCCCTTCAAGAGGC 120
QY 41 H1sAenThrValLyLeuLyVal1ThrLySglYgLyProLeuProPheA1aTPAaP1e 60
Db 121 ACCCAAGCCGCAAGCTGAAGGTGAACAAAGGCGGCGCCCTGCTTCGCTGGAGATC 180
QY 61 LeuSerProGluPheGlu1nYrGlySerLyValYrValLyVSH1sProA1aAP1ePro 80
Db 181 CTGTCCCTCCAGTTCAGATTCAGAGCTTCAAGGCTTCAAGAGCAACCCGCGCATCCTCC 240
QY 81 AaPTrLyLyLeuSerPheProGluGluYPhelySTrPG1uAaGValMetAaPheGlu 100
Db 241 GACTACTTGAAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTGAATGACTTCAG 300
QY 101 AaPG1YgLyValYrVal1ThrVal1ThrGluPheSerSeLyAaNG1uAaPglYCyPhe1eYr 120
Db 301 GAGGGGGGGGTGTGACCGTGACCCAGAGCTCTCTCCGAGAGCGGCGAATTCATCTAC 360
QY 121 LySValLyPhe1LeGlyVal1AaPheProSerAaPglYrProValMetGluLyLyThr 140
Db 361 AAGGTAAAGCTGGCGCGCAACAATTCCTCCGAGCGCCCGCTAATGCAAGAAAGACC 420
QY 141 MetG1YrTrGluAaSer1ThrGluYrGluYrProAaGAPG1VVal1LeuLySglY 160
Db 421 ATGGCGTGGAGGCTTCCACCGAGCGGAATGACCCCGAGAGAGGGCGCTTAAGGGGAG 480
QY 161 11eH1eLyVal1aLeuLyLeuLyAaPglYgLyH1sYrLeuValGluPheLySer1e 180
Db 481 ATCAAGATGAGCTGAAGCTGAAGAGAGCGCGGCCACTAGACCGCGAGTCAAGACACC 540
QY 181 TyrMetA1aLyLyAaProValGluLeuProGlyYrYrYrYrValAaPSeLyLyLeuAaP 200
Db 541 TACATGGCCAAAGAGCCCGGCAAGCTGCGCGCGCTCAAGACCGAGCATCAAGCTGGAC 600
QY 201 11eThSerH1aAaNG1uAaPTr1Thr11eValGluGlu1nYrGluYrThnGluY1AaG 220
Db 601 ATCACTCCCAACAAGAGACTAACAATCTGTGAACAGTAAGAGAGCGCGAGGGCGC 660
QY 221 H1s 221
Db 661 CAC 663

RESULT 14
US-11-218-880-9
/ Sequence 9, Application US/11218880
/ Publication No. US20060003420A1
/ GENERAL INFORMATION:
/ APPLICANT: Tsien, Roger

```

	APPLICANT: Campbell, Robert
/	TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/	TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/	FILE REFERENCE: UC083.1CP2CPI
/	CURRENT APPLICATION NUMBER: US/1/1218, 890
/	PRIOR FILING DATE: 2005-09-01
/	PRIOR APPLICATION NUMBER: US/10/121,258
/	PRIOR FILING DATE: 2002-04-10
/	PRIOR APPLICATION NUMBER: 09/794,308
/	PRIOR FILING DATE: 2001-02-26
/	PRIOR APPLICATION NUMBER: 09/866,538
/	PRIOR FILING DATE: 2001-05-24
/	NUMBER OF SEQ ID NOS: 78
/	SOFTWARE: FastSeq for Windows Version 4.0
/	SEQ ID NO 9
/	LENGTH: 678
/	TYPE: DNA
/	ORGANISM: Artificial Sequence
/	FEATURE:
/	OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
/	US-11-218-880-9
/	"MRFP"
Alignment Scores:	
Pred. No.:	1,36e-110
Score:	1021.00
Percent Similarity:	89.14%
Best Local Similarity:	86.88%
Query Match:	84.10%
DB:	7
Gaps:	0
US-10-006-922A-12 (1-225) x US-11-218-880-9 (1-678)	
QY	1 MetAgsSerSeLyAsnValIllelyGllPhMeKerXphelyValArgMetGluglly 20
DB	1 ATGGCTCTCCCGAGAGGTCAATCAAGAGTTCAATGCGCTTCAAAGTGCCCATGAAGGC 60
QY	21 ThrValAsnGluVHIGluPheGluIllegluGluylglugllyArgProTyrglugly 40
DB	61 TCGGTAAACGCCACAGATTTCAGATCGAAGGCCAAGGCCAAGGCCGCCCTTACGAGGC 120
QY	41 HIsenThrValLysLeuLysValThrllygilyProLeuProPhealattPaaplle 60
DB	121 ACCCAGACCGCAACTGAAGGTGACCAAGGCGGCCCTTGCCTTGCCTTGGAGATC 180
QY	61 LeuSerProGlnPheGluTyrGlyseryValTyrValIlyshISpRoAlaApilPro 80
DB	181 CTGTCCCCCTCAATTCAGTAGCGCTCCAAGGCTTAAGTAAGAACCCCGCGACATCCC 240
QY	81 AsPTyryLysLeuSerPheProGluGlyPheLySTrGluArqValMetAenpheGlu 100
DB	241 GACTAATTMAAGCTGTCTTCCCGAAGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG 300
QY	101 AspGlyGlyValValThrlValThrlGlnAspSerSerLeuGlnAspGlyCySpheIlleTy 120
DB	301 GACGGCGGCGTGTGACCGTGACCAGAGATCTTCCCTGCGAGCGCGAGTTCACTTAC 360
QY	121 LysValLysPheIlleglyValAsnPheProSerAspGlyProValIMetGlnLysThr 140
DB	361 AAGGTAAACCTGGCGGCACAACCTTCCCTTCCAGCGGCCCGCGTAATGCAAGAGAAGCC 420
QY	141 MeGlyTYrGluNlaSerThrGluTrgleutyrrProAspAspGlyValLeuLysGlyLu 160
DB	421 ATGGCTGGAGAGCTCCCAACGAGGANTGATCCCGAGAGCGGCGCGCTGAAGGGGAG 480
QY	161 IlleIleValalaLeuLysLeuLysAspGlyGlyIlyshIlyrLeuValGIupheLysSetlle 180
DB	481 ATCAAGATGAGGCTGAAGCTGAAGGACGCGCGCCACTACGACGCGCGAGTCAACACACC 540
QY	181 TyrMeGlyLysLysProValGlnLeuProGlyTyrrTyrrTyrrValLaaspSerLysLeuAsp 200
DB	541 TCATGGCCAAAGAGCCGATGACGTGCGCGCGCGCTTCAAGACCGGACATCAAGCTGGAC 600

QY 201 IleThrsrHisangluAspTyrThrIleValGluGlnIleGlyArgThrGluGlyArg 220  
DB 601 ATCACTCTCCACAAAGAGACTACACATCGTGAACAGTACGAGCGCGCCGAGCGCCG 660

QY 221 His 221  
DB 661 CAC 663

RESULT 15  
US-10-209-208-80  
Sequence 80, Application US/10209208  
Publication No. US20050244921A1  
GENERAL INFORMATION:  
APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert  
APPLICANT: Geofitrey Baird  
TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
FILE REFERENCE: UC083.1CP2CP2  
CURRENT APPLICATION NUMBER: US/10/209,208  
CURRENT FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 80  
LENGTH: 678  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
OTHER INFORMATION: "mrfp1.1"

US-10-209-208-80  
US-10-006-922A-12 (1-225) x US-10-209-208-80 (1-678)

Alignment Scores:  
Pred. No.: 1,58e-109 Length: 678  
Score: 1012.00 Matches: 190  
Percent Similarity: 88.69% Conservative: 6  
Best Local Similarity: 85.97% Mismatches: 25  
Query Match: 83.36% Indels: 0  
DB: 6 Gaps: 0

QY 1 MetArgSerSerLysAsnValIleIleGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGGCTCTCTCCGAGGAGCTCATCAAGAGTTTCATGCCCTTCAGAGTCCGATGAGGAG 60

QY 21 ThrValangluYHisGluPheGluIleGluGluGluGluGluGluGluGluGluGluGly 40  
DB 61 TCCGTGAACGCGCACAGATTCGAGATCGAGGCGAGGCGAGGCGCGCCCTTCAGAGGAG 120

QY 41 HisAsnThrValLysLeuLysValThrLysGluGlyProLeuProPheAlaTyrAspIle 60  
DB 121 ACCCAGACCGCCCAAGCTGAAGTGAACCAAGGCGCGCCCTTCGCTTCGGAGCATC 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCTCGATTCATACGCTCCCAAGGCTTCAGAGTGAAGCCCGCGGAGCATCCCC 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTACTTGAAGCTGTCTCTCCCGAGGCTTCAGAGTGAAGCCCGTGAATGACTTCGAG 300

QY 101 AspglyGlyValValThrValThrGlnAspSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGCGCGCGTGTGACCGTGAACCAAGCATCTCCCTGCGAGACGCGGAGATTCTATC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

DB 361 AAGGTGAAGCTGGCGCGGACCACTTCCCTCCGACGCGCCCGGATGCAAGAAAGACC 420

QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLleuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCGCTCTCCCGAGCGGATGTACCCCGAGGACGCGCCCTGAAAGGCGAG 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAAGCTGAAGCTGAAGAGCGCGCCACTACAGCCCGAGTCAAGAACACC 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCGGCTTCAAGAACCGATCAAGCTGAGC 600

QY 201 IleThrsrHisangluAspTyrThrIleValGluGlnIleGlyArgThrGluGlyArg 220  
DB 601 ATCACTCTCCACAAAGAGACTACACCATCGTGAACAGTACGAGCGCGCCGAGGCGCCG 660

QY 221 His 221  
DB 661 CAC 663

Search completed: January 12, 2006, 12:46:43  
Job time : 389 secs



98	1085.5	89.4	230	3	AAB01625	AAB01625	Discosoma
99	1085.5	89.4	230	4	AAE00378	AAE00378	Discosoma
100	1085.5	89.4	230	5	ABP70038	ABP70038	Colour Fa
101	1085.5	89.4	230	6	ABP96651	ABP96651	Red Fluor
102	1058	87.1	205	4	AAE00375	AAE00375	Discosoma
103	1054	86.8	205	4	AAE00384	AAE00384	Discosoma
104	1053	86.7	205	4	AAE00383	AAE00383	Discosoma
105	1021	84.1	225	7	ADG24133	ADG24133	Discosoma
106	1021	84.1	225	7	ADL46210	ADL46210	Discosoma
107	1021	84.1	225	8	ADQ59566	ADQ59566	Discosoma
108	1021	84.1	225	9	AEA54910	AEA54910	Fluoresce
109	1012	83.4	225	7	ADL46281	ADL46281	Discosoma

ALIGNMENTS

RESULT 1  
AA99836 standard; protein; 225 AA.

AA99836;  
12-SEP-2003 (revised)  
19-SEP-2000 (first entry)  
Discosoma sp. "red" novel fluorescent protein drfp583.  
Anthozoa; drfp583; fluorescent protein; non-bioluminescent organism;  
fluorescent labeling.  
Discosoma sp; "red".

Key Location/Qualifiers  
Misc-difference 122 /note="encoded by TC"  
Misc-difference 127 /note="encoded by GTTG"  
WO200034326-A1.

15-JUN-2000.  
10-DEC-1999; 99WO-US029473.  
11-DEC-1998; 98US-00210330.  
14-OCT-1999; 99US-00418529.

(CLON-) CLONTECH LAB INC.

Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
Ding L;  
WPI; 2000-423381/36.

Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
useful for fluorescent labeling and as markers.

Claim 20; Page 74-75; 86pp; English.

The present sequence is a novel fluorescent protein (nfp) encoded by the  
full-length cDNA drfp583. drfp583 was isolated from Discosoma sp. "red",  
a non-bioluminescent species of the Class Anthozoa. Fluorescent proteins  
can be used in fluorescent labeling, a useful tool for marking a protein,  
cell or organism of interest. Unlike other markers used in protein  
labeling, such as beta-galactosidase and luciferase, fluorescent proteins  
do not require an exogenous cofactor or substrate. Methods involving  
fluorescent proteins are also less laborious and less difficult to  
control than the traditional methods of fluorescent labeling, where a  
protein of interest is purified and then covalently conjugated to a  
fluorophore derivative. Novel fluorescent proteins isolated from species  
of the Class Anthozoa can be used as markers for gene expression and  
protein localization studies, and in fluorescence resonance energy

transfer (FRET) reactions. They may have improved properties and better  
suitability for larger excitations compared to prior art fluorescent  
proteins such as green fluorescent protein. (Updated on 12-SEP-2003 to  
standardise OS field)

Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRPFKVRMEGTIVNGHFEPIEGEGSRPYEGHNTVTLKVTGKGPLPAMD 60  
DB 1 MRSSKNVKEFRPFKVRMEGTIVNGHFEPIEGEGSRPYEGHNTVTLKVTGKGPLPAMD 60  
QY 61 LSPQFQYGSKYVVKHAPADIPDYKKLSFPEGKRWERNVNFEDGGVTVTVQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVVKHAPADIPDYKKLSFPEGKRWERNVNFEDGGVTVTVQDSSLQDGCFTY 120  
QY 121 KYKFTGVNPPSDGPMQKKTGMEASTERTLYPRDGLKGEIHKALKDGGHYLVFNSI 180  
DB 121 KYKFTGVNPPSDGPMQKKTGMEASTERTLYPRDGLKGEIHKALKDGGHYLVFNSI 180  
QY 181 YWAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
DB 181 YWAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 2  
AAB01622 standard; protein; 225 AA.

AAB01622;  
12-DEC-2000 (first entry)  
Discosoma sp. red fluorescent protein drfp583.

Anthozoa; fluorescent protein; fluorescence labeling; amFP486; cFP484;  
zFP506; zFP538; drfp583; drFP483; asFP600; drfp512; dmFP592.

Discosoma sp.  
WO200034526-A1.  
15-JUN-2000.  
10-DEC-1999; 99WO-US029405.

11-DEC-1998; 98US-00210330.  
(CLON-) CLONTECH LAB INC.

Lukyanov SA, Fradkov AF, Labas YA, Matz MV;  
WPI; 2000-423451/36.

Novel method for identifying a DNA sequence encoding fluorescent proteins  
from non-bioluminescent Anthozoa which are useful for fluorescent  
labeling and as markers.

Claim 3; Page 68-69; 73pp; English.

The present sequence is Discosoma sp. red fluorescent protein drfp583. It  
was isolated using the Aegea victoria green fluorescent protein (GFP)  
sequence, which was used to design PCR primers which might isolate other  
fluorescent proteins from a number of species of Anthozoa. These were  
Anemona majano, Clavularia sp., Zoanthus sp., Discosoma sp. and Anemona  
sulcata. The cDNA obtained was then screened in the search for sequences  
encoding fluorescent proteins. The other proteins found in this manner  
were cFP484, zFP506, zFP538, amFP486, drFP483, asFP600, drfp512 and  
dmFP592. These proteins can be used as fluorescent labels (for gene  
expression and protein localisation studies and in fluorescence resonance



CC energy transfer (FRET) studies) in place of fluorophore derivatives and  
 CC luciferases, as these involve laborious processes and the latter require  
 CC cofactors. They can also be used in place of GFP, which is too stable to  
 CC be useful when studying short-term or repetitive events

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 3; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKVIKEFMRFKRMSTGVNGHEFEIEGEGRPGYGHNTVKLKVTGGGLPFPAMD 60  
 DB 1 MRSSKVIKEFMRFKRMSTGVNGHEFEIEGEGRPGYGHNTVKLKVTGGGLPFPAMD 60  
 QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPGGFKMERVMNFDGGVVTVDSSIQDGCFTY 120  
 DB 61 LSPQFQYGSKVYVKHPADIPDYKLSFPGGFKMERVMNFDGGVVTVDSSIQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVWQKKTMGWEASTERLYPRDGLKGEIHKALKKDGHTLVFEKSI 180  
 DB 121 KYKFIGVNPSPDGPVWQKKTMGWEASTERLYPRDGLKGEIHKALKKDGHTLVFEKSI 180  
 QY 181 YMAKKEVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 DB 181 YMAKKEVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 3  
 AAG65509 standard; protein; 225 AA.

XX AAG65509;

DT 30-NOV-2001 (first entry)

XX Anthozoan red fluorescent protein sequence.

XX Fluorescent protein; Anthozoan; fluorescence; marker; FRET; red.

XX Anthozoa.

XX MO200162919-A1.

XX 30-AUG-2001.

XX 13-FEB-2001; 2001WO-US004625.

XX 23-FEB-2000; 2000US-0184732P.

XX (AURO-) AURORA BIOSCIENCES CORP.

XX Nelson D, Zamatra E, Tsien R;

XX WPI; 2001-557704/62.

XX Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise  
 PT functional red fluorescent proteins, and the encoding nucleic acids, with  
 PT key mutations for improving the proteins function.

XX Disclosure; Page 85; 90pp; English.

XX The invention provides a nucleic acid encoding functional red fluorescent  
 CC protein (II) that differs from the sequence of an Anthozoan red  
 CC fluorescent protein by at least one amino acid substitution, and with  
 CC different fluorescent properties. The red fluorescent protein of the  
 CC invention can be expressed by standard recombinant methodology. (II) are  
 CC used as fluorescent markers and FRET partners. It is used for identifying  
 CC protein-protein interactions. (II) is also suitable for multiplexed  
 CC fluorescent analysis and FRET-based applications using existing Aequorea  
 CC fluorescent proteins. (II) has improved brightness, reduced spectral  
 CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
 CC The key mutations in the encoding nucleic acids provide improved folding,

CC brightness, and create (II) with sharper, more defined excitation and  
 CC emission peaks when expressed in mammalian cells. The present sequence  
 CC represents an anthozoan fluorescent protein

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 4; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKVIKEFMRFKRMSTGVNGHEFEIEGEGRPGYGHNTVKLKVTGGGLPFPAMD 60  
 DB 1 MRSSKVIKEFMRFKRMSTGVNGHEFEIEGEGRPGYGHNTVKLKVTGGGLPFPAMD 60  
 QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPGGFKMERVMNFDGGVVTVDSSIQDGCFTY 120  
 DB 61 LSPQFQYGSKVYVKHPADIPDYKLSFPGGFKMERVMNFDGGVVTVDSSIQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVWQKKTMGWEASTERLYPRDGLKGEIHKALKKDGHTLVFEKSI 180  
 DB 121 KYKFIGVNPSPDGPVWQKKTMGWEASTERLYPRDGLKGEIHKALKKDGHTLVFEKSI 180  
 QY 181 YMAKKEVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 DB 181 YMAKKEVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 4  
 ABB08834 standard; protein; 225 AA.

XX ABB08834;

DT 29-MAY-2002 (first entry)

XX Yeast optimised RFP SEQ ID NO 17.

XX Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;

XX Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;

XX Escherichia coli; green fluorescent protein; biotechnology.

XX Anthozoa.

XX DE20001395-U1.

XX 15-MAR-2001.

XX 27-JAN-2000; 2000DE-02001395.

XX 27-JAN-2000; 2000DE-02001395.

XX (GPCB-) GPC BIOTECH AG.

XX WPI; 2002-228394/29.

XX New DNA encoding red fluorescent protein, useful as marker in  
 PT biotechnology, has sequence optimized for expression in eukaryotes,  
 PT especially yeast or plants.

XX Disclosure; Page 13-14; 19pp; German.

XX The invention relates to DNA (I) containing either sequence ABA95905 or  
 CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
 CC (yRFP). (I) are used to express red fluorescent protein (RFP) in  
 CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
 CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
 CC Arabidopsis thaliana and also in prokaryotes, especially bacteria.  
 CC especially Escherichia coli. RFP is useful in the same way as green  
 CC fluorescent protein but is more generally applicable in modern  
 CC biotechnology. (I) are optimised for expression in yeast and so generate  
 CC RFP at higher levels with stronger fluorescence and thus lowers the  
 CC detection limit and gives a better signal-to-noise ratio. The present

CC sequence is that of the yeast optimised RFP  
XX SQ Sequence 225 AA;  
Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSSKNVIEKPEFRFKRMGEVTNGHEFEIEGEGEGRPYEGHNTVTKLTKTGGLPFPAMD 60  
DB 1 MSSKNVIEKPEFRFKRMGEVTNGHEFEIEGEGEGRPYEGHNTVTKLTKTGGLPFPAMD 60  
QY 61 LSPQFOYSGSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYSGSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIVGNFSPDGPVMQKKTWGEASTERLYPRDGLKGEIHKALKDKGHIYVEFKSI 180  
DB 121 KVKFIVGNFSPDGPVMQKKTWGEASTERLYPRDGLKGEIHKALKDKGHIYVEFKSI 180  
QY 181 YMAKRPVOLPGYIYVDSKLDITSHNEDYTIIVEQYERTGRRHFL 225  
DB 181 YMAKRPVOLPGYIYVDSKLDITSHNEDYTIIVEQYERTGRRHFL 225  
RESULT 5  
AAE28833  
ID AAE28833 standard; protein; 225 AA.  
AC AAE28833;  
XX  
DT 27-DEC-2002 (first entry)  
XX  
DB Discosoma sp. drp583 (NFP-6) wild-type protein.  
XX  
KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
fluorescence activated cell sorting application; fluorescent timer;  
biosensor; fluorescence resonance energy transfer application; FRET;  
colouring agent; recombinant DNA application; analyte detection assay;  
buncscreen; second messenger detector; drp583 protein; NFP-6.  
XX  
OS Discosoma sp.  
XX  
PN WO200268459-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 20-FEB-2002; 2002WO-US005749.  
XX  
PR 21-FEB-2001; 2001US-0270983P.  
XX  
PR 04-DEC-2001; 2001US-00006922.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
XX  
DR MPI; 2002-691654/74.  
XX  
DR N-PSDB; AAD46278.  
XX  
PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
analyte detection assays or fluorescence activated cell sorting  
applications.  
XX  
PS Disclosure; Page 70-71; 80pp; English.  
XX  
CC The invention relates to nucleic acid molecules encoding non-aggregating  
chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as buncscreens or filters, in fluorescence  
resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
messenger detectors, in fluorescence activated cell sorting applications,

CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drp583 (NFP-6) wild-type protein of the  
CC invention  
XX SQ Sequence 225 AA;  
Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSSKNVIEKPEFRFKRMGEVTNGHEFEIEGEGEGRPYEGHNTVTKLTKTGGLPFPAMD 60  
DB 1 MSSKNVIEKPEFRFKRMGEVTNGHEFEIEGEGEGRPYEGHNTVTKLTKTGGLPFPAMD 60  
QY 61 LSPQFOYSGSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYSGSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIVGNFSPDGPVMQKKTWGEASTERLYPRDGLKGEIHKALKDKGHIYVEFKSI 180  
DB 121 KVKFIVGNFSPDGPVMQKKTWGEASTERLYPRDGLKGEIHKALKDKGHIYVEFKSI 180  
QY 181 YMAKRPVOLPGYIYVDSKLDITSHNEDYTIIVEQYERTGRRHFL 225  
DB 181 YMAKRPVOLPGYIYVDSKLDITSHNEDYTIIVEQYERTGRRHFL 225  
RESULT 6  
AAE17540  
ID AAE17540 standard; protein; 225 AA.  
AC AAE17540;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DB Discosoma sp. humanised wild-type Anthozoa protein drp583.  
XX  
KW Fluorescent timer protein; protein movement; translocation; trafficking;  
promoter activity; gene expression; transgenic plant; gene modification;  
protein age; anthozoa protein; drp583.  
XX  
OS Discosoma sp.  
XX  
PN WO200196373-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019097.  
XX  
PR 14-JUN-2000; 2000US-0211607P.  
XX  
PR (CLON-) CLONTECH LAB INC.  
XX  
PI Fradkov AF, Terekikh A;  
XX  
DR MPI; 2002-154595/20.  
XX  
DR N-PSDB; AAD28207.  
XX  
PT New fluorescent timer proteins comprising an emission spectrum that  
changes over time from a first wavelength to a second wavelength, useful  
for monitoring intracellular protein movement, translocation, trafficking  
or stability.  
XX  
PS Example 1; Fig 1; 89pp; English.  
XX  
CC The invention relates to a fluorescent timer protein having an emission  
CC spectrum that changes over time after synthesis from a first wavelength  
CC to a second wavelength. The fluorescent timer proteins are useful in  
CC monitoring the activity of a promoter, determining the age of a protein,  
CC identifying an agent that modulates the activity of a promoter and in  
CC enriching a population of cells comprising a fluorescent timer protein.  
CC The fluorescent timer proteins are also useful for assessing gene  
expression during development of a multicellular organism or during

cellular differentiation, in response to a drug or other inducer of promoter activity, as a reporter to serve as a read-out of promoter activity, monitoring intracellular protein movement or translocation, protein trafficking, or protein stability, to investigate temporal aspects of the activity of a regulatory element, for determining cell fate during development and organ remodeling, in spatial and temporal visualization of newly synthesised proteins and accumulated proteins, e.g. in distinguishing between newly formed and pre-existing structures, e.g. membrane junctions and extracellular matrix components. The fluorescent timer proteins may further be used to investigate where photobleaching techniques are employed, as detectable labels, as selectable markers, as biosensors in prokaryotic and eukaryotic cells, in protease cleavage assays, and as second messenger detectors. The nucleic acids can be used to generate transgenic, non-human plants or animals or site-specific gene modifications in cell lines. The present sequence is *Discosoma* sp. fluorescent wild-type Anthozoa protein drps83 used for generating fluorescent proteins

Sequence 225 AA:

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSSSKNVIKKFMKFKVMKGTGNGHFEIIEGEGGAPRYEGHNTVTLKVTKGGLPFPAMD 60  
1 MSSSKNVIKKFMKFKVMKGTGNGHFEIIEGEGGAPRYEGHNTVTLKVTKGGLPFPAMD 60  
61 LSPQFQYSGKVVYKHPADIPDYKLSFPEGFKMERVMNPFEDGAVVTQDSSLQDGCFTY 120  
61 LSPQFQYSGKVVYKHPADIPDYKLSFPEGFKMERVMNPFEDGAVVTQDSSLQDGCFTY 120  
121 KYKFIGVNFPSDGPVWQKKTGMEASTERLYPRDGYLKEIHKALKDGGHYLVEFKSI 180  
121 KYKFIGVNFPSDGPVWQKKTGMEASTERLYPRDGYLKEIHKALKDGGHYLVEFKSI 180  
181 YMAKKRVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
181 YMAKKRVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 7

AAO18270 standard; protein; 225 AA.

AAO18270;

26-SBP-2002 (first entry)

*Discosoma* red fluorescent protein.

Yeast; RAD51; promoter; genotoxicity cassette; cytotoxicity cassette; modified yeast strain; environmental pollution.

*Discosoma* sp.

DE10061872-A1.

20-JUN-2002.

12-DEC-2000; 2000DB-01061872.

12-DEC-2000; 2000DB-01061872.

(LICH/) LICHTENBERG-FRATTE H.

Lichtenberg-Fratt H;

WPI; 2002-539633/58.

N-PSDB; AAL47952.

Modified yeast strain, useful for detecting toxic compounds in environment, contains integrated cassettes responsive to genotoxic and

cytotoxic compounds.

Disclosure; Page 21-22; 34pp; German.

The present invention relates to a modified yeast strain that contains, integrated stably and functionally in its genome, a genotoxicity cassette and a cytotoxicity cassette, each comprising a promoter and reporter gene, both of which are different in the two cassettes. The modified yeast strain is used to detect environmental pollution, especially genotoxic and/or cytotoxic substances in complex environmental contaminants, especially organic compounds, but also (non-)ionising radiation (waste) water (e.g. as an early warning system), medical monitoring (waste) water and for industrial process control. The present toxicology screening and for industrial process control. The present sequence is a marker protein suitable for use in the cassettes of the present invention

Sequence 225 AA:

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSSSKNVIKKFMKFKVMKGTGNGHFEIIEGEGGAPRYEGHNTVTLKVTKGGLPFPAMD 60  
1 MSSSKNVIKKFMKFKVMKGTGNGHFEIIEGEGGAPRYEGHNTVTLKVTKGGLPFPAMD 60  
61 LSPQFQYSGKVVYKHPADIPDYKLSFPEGFKMERVMNPFEDGAVVTQDSSLQDGCFTY 120  
61 LSPQFQYSGKVVYKHPADIPDYKLSFPEGFKMERVMNPFEDGAVVTQDSSLQDGCFTY 120  
121 KYKFIGVNFPSDGPVWQKKTGMEASTERLYPRDGYLKEIHKALKDGGHYLVEFKSI 180  
121 KYKFIGVNFPSDGPVWQKKTGMEASTERLYPRDGYLKEIHKALKDGGHYLVEFKSI 180  
181 YMAKKRVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
181 YMAKKRVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 8

ADY51716 standard; protein; 225 AA.

ADY51716;

05-MAY-2005 (first entry)

*Discosoma* sp. wild type red fluorescent protein (RFP) Seq 12.

fluorescence; mutagenesis; red fluorescent protein; protein interaction.

*Discosoma* sp.

WO200268605-A2.

06-SBP-2002.

26-FEB-2002; 2002MO-US006063.

26-FEB-2001; 2001US-00794308.

24-MAY-2001; 2001US-00866538.

(REGC ) UNIV CALIFORNIA.

Tsien RY, Baird GS, Campbell RE, Zacharias DA;

WPI; 2002-713372/77.

N-PSDB; ADY51715.

New non-oligomerizing fluorescent protein containing at least one mutation that reduces or eliminates the ability of the protein to oligomerize, useful for making better and new assays for molecular

PT biology.  
XX  
PS Claim 10; SEQ ID NO 12; 117bp; English.  
XX  
CC This invention relates to a novel non-oligomerizing fluorescent protein.  
CC Specifically, it refers to the presence of at least one mutation in the  
CC fluorescent protein that reduces or eliminates the ability of the protein  
CC to oligomerize. The present invention describes fluorescent proteins and  
CC derived from naturally occurring green or red fluorescent proteins and  
CC provides a fusion protein that comprises a non-oligomerizing fluorescent  
CC protein linked to at least one protein of interest. As such, these fusion  
CC proteins can be used in methods and compositions to determine the pH of a  
CC sample, or whether the sample contains an enzyme, molecule or agent that  
CC regulates the activity of an expression control sequence. Furthermore,  
CC they may be used to identify a specific interaction of molecules, such  
CC that they are useful for improving or developing new assays in the field  
CC of molecular biology. This polypeptide sequence is the Discosoma sp. wild  
CC type red fluorescent protein (RFP) of the invention.  
XX  
SQ Sequence 225 AA;  
Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLTKTGGPLPFANDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLTKTGGPLPFANDI 60  
QY 61 LSPQFQYGSKYVYVHPADIPDYKKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVYVHPADIPDYKKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
QY 121 KVFEGVNPSPSGPVMQKKTGMWEASTERYLPDGLVKGELHKALKDGGHYLVFPRSI 180  
DB 121 KVFEGVNPSPSGPVMQKKTGMWEASTERYLPDGLVKGELHKALKDGGHYLVFPRSI 180  
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTYVEQYERTBGRHHLFL 225  
DB 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTYVEQYERTBGRHHLFL 225  
RESULT 9  
AAE34962  
ID AAE34962 standard; protein; 225 AA.  
XX  
AC AAE34962;  
XX  
DT 28-MAY-2003 (first entry)  
XX  
DE Discosoma species red fluorescent protein (RFP).  
XX  
KW Phosphorylation indicator; fluorescent protein; detection; phosphatase;  
KM kinase; red fluorescent protein; RFP.  
XX  
OS Discosoma sp.  
XX  
PN WO200295058-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 24-MAY-2002; 2002WO-US016955.  
XX  
PR 24-MAY-2001; 2001US-00865291.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Tsien RY, Ting AY, Zhang J;  
XX  
DR WPI: 2003-148474/14.  
XX  
PT N-PSDB; AAD53432.  
PT Novel chimeric phosphorylation indicators, useful for detecting

PT kinase/phosphatase in samples, has donor molecule, phosphorylatable  
PT domain, phosphoaminoacid binding domain, and acceptor molecule, in  
PT operative linkage.  
XX  
PS Disclosure; Col 65-66; 38pp; English.  
XX  
CC The present invention relates to chimeric phosphorylation indicators  
CC comprising a phosphorylation polypeptide and a fluorescent protein or in  
CC operative linkage, a donor molecule, a phosphorylatable domain, a  
CC phosphoaminoacid binding domain (PABD) and an acceptor molecule. The  
CC phosphorylation indicators of the invention are useful for detecting  
CC kinases or phosphatases in a biological sample. They are also useful in  
CC high throughput analysis e.g. for detecting a kinase inhibitor or  
CC phosphatase inhibitor. The present sequence is Discosoma species red  
CC fluorescent protein (RFP) used in the invention  
XX  
SQ Sequence 225 AA;  
Query Match 100.0%; Score 1214; DB 6; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLTKTGGPLPFANDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLTKTGGPLPFANDI 60  
QY 61 LSPQFQYGSKYVYVHPADIPDYKKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVYVHPADIPDYKKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
QY 121 KVFEGVNPSPSGPVMQKKTGMWEASTERYLPDGLVKGELHKALKDGGHYLVFPRSI 180  
DB 121 KVFEGVNPSPSGPVMQKKTGMWEASTERYLPDGLVKGELHKALKDGGHYLVFPRSI 180  
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTYVEQYERTBGRHHLFL 225  
DB 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTYVEQYERTBGRHHLFL 225  
RESULT 10  
ADC24126  
ID ADC24126 standard; protein; 225 AA.  
XX  
AC ADC24126;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Discosoma wild-type red fluorescent protein.  
XX  
DE Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
KM fluorescent protein variant; transcription induction detection;  
KW fluorescence energy resonance transfer; FRET; protein kinase;  
KM protein phosphatase; ion indicator.  
XX  
OS Discosoma.  
XX  
PN US2003059835-A1.  
XX  
PD 27-MAR-2003.  
XX  
PF 10-APR-2002; 2002US-00121258.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
XX  
PR 24-MAY-2001; 2001US-00866538.  
XX  
PA (TSIE/) TSIE R Y.  
XX  
PA (CAMP/) CAMPBELL R E.  
XX  
PI Tsien RY, Campbell RE;  
XX  
DR WPI: 2003-743764/70.  
XX  
PT N-PSDB; ADC24127, ADC24134.

PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.

PS Claim 1; SEQ ID NO 1; 67bp; English.

CC The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (Dared) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type Dared sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the Dared variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cell  
CC containing a vector which comprises (i) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localization or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcriptions, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as  $Ca^{2+}$ ,  $Zn^{2+}$ , for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This is the amino acid sequence of  
CC Discosoma wild-type red fluorescent protein.

CC Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIEKFMKFKVMKMGTVNGHFEIIEGEGRPYEGHNTVTKLKTGKGPLPAMDI 60  
DB 1 MRSSKNVIEKFMKFKVMKMGTVNGHFEIIEGEGRPYEGHNTVTKLKTGKGPLPAMDI 60  
QY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNEDGCVTVTODSSLQDSCFTY 120  
DB 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNEDGCVTVTODSSLQDSCFTY 120  
QY 121 KYKFIQVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGCHYLVEFKSI 180  
DB 121 KYKFIQVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGCHYLVEFKSI 180  
QY 181 YMAKKRVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
DB 181 YMAKKRVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 11

ID ABW00918 standard; protein; 225 AA.

AC ABW00918;

DT 15-JAN-2004 (first entry)

DB Discosoma sp. red fluorescent protein (RFP).

KM Fluorescent protein; resonance energy transfer; pH; detection;

KW red fluorescent protein; RFP.

OS Discosoma sp.

XX US2003170911-A1.

XX 11-SEP-2003.

XX 26-FEB-2001; 2001US-00794308.

XX 26-FEB-2001; 2001US-00794308.

XX (TSIE//) TSIE R Y.

XX (ZACH//) ZACHARIAS D A.

XX (BAIR//) BAIRD G S.

XX Tsien RY, Zacharias DA, Baird GS;

XX WPI; 2003-802418/75.

XX N-PSDB; AAD61969.

PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.

PS Claim 10; Page 30-31; Opp; English.

CC The invention relates to a non-oligomerizing fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerize. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP)

CC Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIEKFMKFKVMKMGTVNGHFEIIEGEGRPYEGHNTVTKLKTGKGPLPAMDI 60  
DB 1 MRSSKNVIEKFMKFKVMKMGTVNGHFEIIEGEGRPYEGHNTVTKLKTGKGPLPAMDI 60  
QY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNEDGCVTVTODSSLQDSCFTY 120  
DB 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNEDGCVTVTODSSLQDSCFTY 120  
QY 121 KYKFIQVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGCHYLVEFKSI 180  
DB 121 KYKFIQVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGCHYLVEFKSI 180  
QY 181 YMAKKRVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
DB 181 YMAKKRVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 12

ID ADF70403 standard; protein; 225 AA.

AC ADF70403;

DT 12-FEB-2004 (first entry)

DB Discosoma wild-type GFP variant protein SeqID26.

KM ligand; orphan receptor protein; fusion protein; fluorescent protein;

KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;

KM GFPuv; Enhanced GFP; EGFP.

OS Discosoma sp.

XX WO2003071272-A1.

XX 28-AUG-2003.  
 PD  
 XX  
 PF 21-FEB-2003; 2003WO-JP001901.  
 XX  
 PR 22-FEB-2002; 2002JP-00045728.  
 XX  
 PR 23-JUL-2003; 2002JP-00213949.  
 XX  
 PR 11-OCT-2002; 2002JP-00298237.  
 XX  
 (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PA Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;  
 XX  
 PI WPI; 2003-697654/66.  
 XX  
 DR N-PSDB; ADF70404.  
 XX  
 PT Transformation of cells with a fusion protein of an orphan receptor  
 PT protein with a fluorescent protein useful for identification of ligands  
 PT to the orphan receptor.  
 XX  
 PS Disclosure; SEQ ID NO 26; 594pp; Japanese.  
 XX  
 CC This invention relates to a novel method of identifying ligands to an  
 CC orphan receptor protein which comprises transforming cells with DNA  
 CC encoding a fusion protein of the orphan receptor with a fluorescent  
 CC protein, so that the fusion protein is expressed in the cells (or cell  
 CC membranes isolated from them) and contacting the cells with the potential  
 CC ligand to be tested. A suitable fluorescent protein for incorporation in  
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
 CC identification of ligands binding to an orphan receptor protein.  
 XX  
 SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1e-127;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSKNVITKEPMPFKYRMGCTNGHGFETEGSGEGSPGYCHNTVTKYTKGGPLPRAWMI 60  
 DB 1 MSSSKNVITKEPMPFKYRMGCTNGHGFETEGSGEGSPGYCHNTVTKYTKGGPLPRAWMI 60  
 QY 61 LSPFOYGSKVVYKHPADIPDYKLSFPPGFKMERVMNFDGAVVTYVTDSSLDQGCFTY 120  
 DB 61 LSPFOYGSKVVYKHPADIPDYKLSFPPGFKMERVMNFDGAVVTYVTDSSLDQGCFTY 120  
 QY 121 KYKFTGVNFPSPGSPVMOQKKTGMEASTERTLYPRDGYLKGEIHKALKKDGGHYLVEPKSI 180  
 DB 121 KYKFTGVNFPSPGSPVMOQKKTGMEASTERTLYPRDGYLKGEIHKALKKDGGHYLVEPKSI 180  
 QY 181 YNAKKPVOLPGYVYVDSKLDITSHNDYITVBOYRTEGRHHLFL 225  
 DB 181 YNAKKPVOLPGYVYVDSKLDITSHNDYITVBOYRTEGRHHLFL 225  
 ID ADH34489 standard; protein; 225 AA.  
 AC ADH34489;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Diacosoma sp. red fluorescent protein Dared (wild-type).  
 XX  
 XX Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;  
 KW Chitnarian; Anthozoan; labelling; colouring agents; pigment;  
 KW analyte detection assay; selectable marker; sunscreen; selective filter;  
 KW fluorescence resonance energy transfer; FRET; biosensor;  
 KW whole cell marker; second messenger detector; in vivo marker;  
 KW fluorescence activated cell sorting; fluorescent timer;  
 KW red fluorescent protein; Dared.  
 XX

OS Diacosoma sp.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 146 /note= "Key position for chromoprotein/fluorescent  
 FT protein activity. Corresponds to GFP residue 148"  
 FT Misc-difference 161 /note= "Key position for chromoprotein/fluorescent  
 FT protein activity. Corresponds to GFP residue 165"  
 FT Misc-difference 163 /note= "Key position for chromoprotein/fluorescent  
 FT protein activity. Corresponds to GFP residue 167"  
 FT Misc-difference 197 /note= "Key position for chromoprotein/fluorescent  
 FT protein activity. Corresponds to GFP residue 203"  
 XX  
 XX WO2003057833-A2.  
 XX  
 XX 17-JUL-2003.  
 XX  
 XX 23-DEC-2002; 2002WO-US041418.  
 XX  
 XX 26-DEC-2001; 2001US-0343128P.  
 XX  
 XX (CLON-) CLONTECH LAB INC.  
 XX  
 XX Bulina ME, Chudakov D, Lukyanov KA;  
 XX  
 XX WPI; 2003-607998/57.  
 XX  
 XX Novel nucleic acid encoding interconverted mutant of chromo-or  
 XX fluorescent protein which are useful as biosensors, coloring agents.  
 XX  
 XX Example 1; Fig 1; 56pp; English.  
 XX  
 XX The invention relates to interconverted mutants of chromoproteins (CP) or  
 XX fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
 XX derived from a Chitnarian species, preferably a non-bioluminescent  
 XX Chitnarian species, and most preferably an Anthozoan species. The  
 XX invention is based on the finding that although green fluorescent protein  
 XX (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
 XX homology, there are certain positions (referred to as 148, 165, 167 and  
 XX 203; numbering corresponds to GFP) that are occupied by noticeably  
 XX different residues in the two types of proteins. Mutagenesis of the  
 XX residues in these key positions in, for example, a fluorescent protein,  
 XX to those found in a chromoprotein is therefore proposed to confer  
 XX chromoprotein activity on the fluorescent protein mutant, with  
 XX chromoproteins being able to be converted into fluorescent proteins in a  
 XX similar manner. The invention also relates to expression constructs,  
 XX vectors, host cells and host cell progeny comprising a nucleic acid of  
 XX the invention; the recombinant production of an interconverted  
 XX chromoprotein or fluorescent protein mutant; and antibodies specific for  
 XX interconverted mutant proteins of the invention. The interconverted  
 XX fluorescent protein, fluorescent protein mutants having chromoprotein  
 XX activity can be useful as colouring agents in, for example, food  
 XX compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
 XX with chromoprotein activity are also useful as labels in biological  
 XX analyte detection assays, as selectable markers in recombinant DNA  
 XX applications (e.g. the production of transgenic cells and organisms), and  
 XX are also useful as sunscreens and selective filters. Chromoprotein  
 XX mutants having fluorescent protein activity useful in fluorescence  
 XX resonance energy transfer (FRET) applications, as biosensors in  
 XX prokaryotic and eukaryotic cells, as markers of whole cells to detect  
 XX changes in multicellular reorganisation and migration, as second  
 XX messenger detectors, as in vivo markers in animals (e.g., transgenic  
 XX animals), in fluorescence activated cell sorting applications, in  
 XX protease cleavage assays, and in assays to determine the phospholipid  
 XX composition in biological membranes. Proteins with fluorescent protein  
 XX activity can also be used as fluorescent timers, where the switch of one  
 XX fluorescent colour to another (e.g., green to red) is concomitant with  
 XX the ageing of the protein and is useful for determination of the  
 XX activation or deactivation of gene expression. The present sequence

CC represents a wild-type red fluorescent protein, DsRed, from *Discosoma* sp.  
CC that was used as a parent sequence for the generation of mutant proteins  
CC in an example of the invention.  
XX

Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRFRVMEGTNGHEPTEEGEGRPYEGHNTVYLKVTKGSPLPFADI 60  
1 MRSSKNVKEFRFRVMEGTNGHEPTEEGEGRPYEGHNTVYLKVTKGSPLPFADI 60  
DB 1 MRSSKNVKEFRFRVMEGTNGHEPTEEGEGRPYEGHNTVYLKVTKGSPLPFADI 60  
QY 61 LSPQFQYSKYVVKHPADIPYKLSPEEGFKMERVMNFDGVTVTQDSSLQDGCFTY 120  
61 LSPQFQYSKYVVKHPADIPYKLSPEEGFKMERVMNFDGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYSKYVVKHPADIPYKLSPEEGFKMERVMNFDGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKKTWGMESTERLYPRDGLKGEIHKALKDKGGHYLVFEXSI 180  
121 KVKFIGVNFPSDGPVMQKKTWGMESTERLYPRDGLKGEIHKALKDKGGHYLVFEXSI 180  
DB 121 KVKFIGVNFPSDGPVMQKKTWGMESTERLYPRDGLKGEIHKALKDKGGHYLVFEXSI 180  
QY 181 YMAKRPVOLPGYTYVDSKLDITSHNEDYTIIVEOYERTGRHHLFL 225  
181 YMAKRPVOLPGYTYVDSKLDITSHNEDYTIIVEOYERTGRHHLFL 225  
DB 181 YMAKRPVOLPGYTYVDSKLDITSHNEDYTIIVEOYERTGRHHLFL 225

#### RESULT 14

ADL46203  
ID ADL46203 standard; protein; 225 AA.

AC ADL46203;

DT 20-MAY-2004 (first entry)

XX *Discosoma* red fluorescent (DsRed) protein.

XX red fluorescent protein; DsRed; fluorescence; red wavelength;  
XX oligomerization; tetramerization; immunoassay; hybridization assay.

OS *Discosoma* sp.

PN WO2003086446-A1.

PD 23-OCT-2003.

PF 09-APR-2003; 2003WO-US010879.

PR 10-APR-2002; 2002US-00121258.

PR 29-JUL-2002; 2002US-00209208.

XX (REGC ) UNIV CALIFORNIA.

PI Telen RX, Campbell RE, Baird GS;

DR WPI; 2003-845265/78.

DR N-PSDB; ADL46204.

PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.

PS Claim 1; SEQ ID NO 1; 166pp; English.

XX The invention relates to a polynucleotide sequence encoding a *Discosoma*  
XX red fluorescent protein (DsRed) variant having a reduced propensity to  
XX oligomerize. The protein variant comprises one or more amino acid  
XX substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
XX sequence, where the substitutions result in reduced propensity of the  
XX DsRed variant to form tetramers and where the variant displays detectable  
XX fluorescence of at least one red wavelength. The composition and methods  
XX are useful in producing red fluorescent proteins having reduced

CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC protein.  
XX

Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRFRVMEGTNGHEPTEEGEGRPYEGHNTVYLKVTKGSPLPFADI 60  
1 MRSSKNVKEFRFRVMEGTNGHEPTEEGEGRPYEGHNTVYLKVTKGSPLPFADI 60  
DB 1 MRSSKNVKEFRFRVMEGTNGHEPTEEGEGRPYEGHNTVYLKVTKGSPLPFADI 60  
QY 61 LSPQFQYSKYVVKHPADIPYKLSPEEGFKMERVMNFDGVTVTQDSSLQDGCFTY 120  
61 LSPQFQYSKYVVKHPADIPYKLSPEEGFKMERVMNFDGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYSKYVVKHPADIPYKLSPEEGFKMERVMNFDGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKKTWGMESTERLYPRDGLKGEIHKALKDKGGHYLVFEXSI 180  
121 KVKFIGVNFPSDGPVMQKKTWGMESTERLYPRDGLKGEIHKALKDKGGHYLVFEXSI 180  
DB 121 KVKFIGVNFPSDGPVMQKKTWGMESTERLYPRDGLKGEIHKALKDKGGHYLVFEXSI 180  
QY 181 YMAKRPVOLPGYTYVDSKLDITSHNEDYTIIVEOYERTGRHHLFL 225  
181 YMAKRPVOLPGYTYVDSKLDITSHNEDYTIIVEOYERTGRHHLFL 225  
DB 181 YMAKRPVOLPGYTYVDSKLDITSHNEDYTIIVEOYERTGRHHLFL 225

#### RESULT 15

ADN33979  
ID ADN33979 standard; protein; 225 AA.

AC ADN33979;

DT 01-JUL-2004 (first entry)

XX wild-type DsRed protein.

XX *Chlidarian*; fluorescence resonance energy transfer; FRET; wild-type DsRed.

OS *Discosoma* sp.

PN WO2003054158-A2.

PD 03-JUL-2003.

PF 18-DEC-2002; 2002WO-US040539.

PR 19-DEC-2001; 2001US-0341723P.

XX (UYCH-) UNIV CHICAGO.

PI Bevis B, Glick B;

DR WPI; 2003-569236/53.

DR N-PSDB; ADN33978.

PT Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent  
PT mutant of a *Chlidarian* chromo- or fluorescent protein or its mutant,  
PT useful for applications involving chromo- or fluorescent proteins.  
XX

PS Claim 8; SEQ ID NO 2; 65pp; English.

XX The present invention relates to nucleic acid that encodes a rapidly  
XX maturing chromo- or fluorescent mutant of a *Chlidarian* chromo- or  
XX fluorescent protein or its mutant. The protein is useful in applications  
XX involving nucleic acid encoding a chromo- or fluorescent protein which involves  
XX growing the cell, whereby the protein is expressed, and isolating the  
XX protein substantially free of other proteins. The protein is useful in  
XX applications involving chromo- or fluorescent protein and is useful as  
XX PCR primers, hybridization probes, etc. The expression cassettes are

CC useful for synthesizing related proteins. The chromoproteins are useful  
CC as coloring agents which are capable of imparting color or pigment to a  
CC particular composition of matter e.g. food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g., animals and plants. The chromoproteins  
CC may also find use as labels in analytical detection assays, e.g. assays for  
CC biological analyses of interest and as selectable markers in recombinant  
CC DNA applications, e.g. the production of transgenic cells and organisms.  
CC The fluorescent proteins find use in a variety of different applications,  
CC e.g. in fluorescence resonance energy transfer (FRET) applications, as  
CC biosensors in prokaryotic and eukaryotic cells, in applications involving  
CC the automated screening of arrays of cells expressing fluorescent  
CC reporting groups by using microscopic imaging and electronic analysis, as  
CC second messenger detectors, and in fluorescence activated cell sorting  
CC applications and as in vivo marker in animals. The fluorescent proteins  
CC also find use in protease cleavage assays. The proteins can also be used  
CC in assays to determine the phospholipid composition in biological  
CC membranes and as a fluorescent timer. The present sequence represents the  
CC wild-type DsRed.

CC  
XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEPMRPFKRMGEVTNGHEFEIEGEGGRPYEGHNTVTKKTKGGPLPFAWDI 60  
DB 1 MRSSKNVKEPMRPFKRMGEVTNGHEFEIEGEGGRPYEGHNTVTKKTKGGPLPFAWDI 60  
QY 61 LSPQFQYGSKYVYKHPADIPIYKKLSFPEGFKMERVMNFPEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVYKHPADIPIYKKLSFPEGFKMERVMNFPEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGPVWQKKTGMEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
DB 121 KYKFIGVNPSPDGPVWQKKTGMEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
DB 181 YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 16

AD136421  
ID AD136421 standard; protein; 225 AA.

AC AD136421;

DT 22-APR-2004 (first entry)

XX Discosoma sp. red fluorescent protein (RED).

XX Fluorobody; binding ligand; green fluorescent protein; GFP;

KW target detection; red fluorescent protein; RED.

XX Discosoma sp.

OS  
XX  
FH Key Location/Qualifiers

FT 22..26

FT /note= "Loop region"

FT 37..39

FT /note= "Loop region"

FT 75..81

FT /note= "Loop region"

FT 100..103

FT /note= "Loop region"

FT 114..117

FT /note= "Loop region"

FT 153..156

FT /note= "Loop region"

FT 167..170

FT /note= "Loop region"

FT 185..192

FT /note= "Loop region"  
FT 204..209  
FT /note= "Loop region"

PN US2003203355-A1.

PD 30-OCT-2003.

XX 24-APR-2002; 2002US-00132067.

XX 24-APR-2002; 2002US-00132067.

XX 24-APR-2002; 2002US-00132067.

XX (LALA-) LOS ALAMOS NAT LAB.

XX (REGC ) UNIV CALIFORNIA.

XX Bradbury AM, Zeytun A, Waldo GS;

XX WPI, 2004-154325/15.

XX N-PSDB; AD136420.

XX Novel binding ligand with intrinsic fluorescence and comprising

XX detecting target molecule.

XX Claim 2; SEQ ID NO 4; 23pp; English.

XX The invention relates to binding ligands (fluorobodies) with intrinsic

XX fluorescence, which comprises green fluorescent protein (GFP) having

XX heterologous binding sites. The binding ligand is useful for detecting

XX the target molecule and is efficiently detects the target molecule. The

XX present sequence is Discosoma sp. red fluorescent protein (RED) used in

XX the exemplification of the invention.

XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 8; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEPMRPFKRMGEVTNGHEFEIEGEGGRPYEGHNTVTKKTKGGPLPFAWDI 60  
DB 1 MRSSKNVKEPMRPFKRMGEVTNGHEFEIEGEGGRPYEGHNTVTKKTKGGPLPFAWDI 60  
QY 61 LSPQFQYGSKYVYKHPADIPIYKKLSFPEGFKMERVMNFPEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVYKHPADIPIYKKLSFPEGFKMERVMNFPEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGPVWQKKTGMEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
DB 121 KYKFIGVNPSPDGPVWQKKTGMEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
DB 181 YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 17

ADN97769  
ID ADN97769 standard; protein; 225 AA.

AC ADN97769;

DT 01-JUL-2004 (first entry)

XX D sp red fluorescent protein SEQ ID NO: 22.

XX enzyme; sensor cell; green fluorescent protein;

KW signal transduction detection system; promoter; targeting sequence;

XX targeted drug.

XX Discosoma sp.

XX WO2004031415-A2.



XX 15-APR-2004.  
PD 05-SEP-2003; 2003WO-US028078.  
XX 05-SEP-2003; 2002US-0408297P.  
XX 05-SEP-2002; 2002US-0408297P.  
XX (VERT-) VERTEX PHARM INC.  
XX Whitney MA, Zeh K, Sanders PS;  
XX WPI; 2004-330208/30.  
XX N-PSDB; ADM97768.  
XX  
XX Developing a sensor cell, useful in determining the activity of a target  
XX gene and in developing therapeutic drugs, comprises providing cells  
XX comprising a signal transduction detection system and introducing DNA  
XX construct into cells.  
XX  
XX Disclosure; Page 168-169; 234pp; English.  
XX  
XX The present invention relates to a method of developing a sensor cell,  
XX for determining the activity of a target gene in the cell, which  
XX comprises providing a homogeneous population of cells, where each of the  
XX cells comprises a signal transduction detection system and introducing  
XX into the population of cells an isolated DNA construct comprising a  
XX promoter operatively linked to a targeting sequence. The method is useful  
XX in developing a sensor cell for determining the activity of a target gene  
XX in the cell. The sensor cell and the methods are useful in developing new  
XX and therapeutic drugs directed to the targets. The present sequence is a  
XX polypeptide shown in the exemplification of the invention.  
XX  
XX Sequence 225 AA;  
SQ

Query Match 100.0%; Score 1214; DB 8; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSKKNVKEFMFKRMKMEGTNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60  
DB 1 MRSKKNVKEFMFKRMKMEGTNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60  
QY 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120  
QY 121 KVFIFGVNFPSPDGPVWQKKTMGWEASTERLYPRDGYLKGSIHKALKKDGCHYLVEFKSI 180  
DB 121 KVFIFGVNFPSPDGPVWQKKTMGWEASTERLYPRDGYLKGSIHKALKKDGCHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 18  
ADQ28780 standard; protein; 225 AA.  
ID ADQ28780  
XX  
XX ADQ28780;  
AC  
XX  
XX 07-OCT-2004 (first entry)  
DT  
XX  
XX Jellyfish green fluorescent protein DsRed.  
DE  
XX  
XX Green fluorescent protein; fluorescent protein.  
XX  
XX Aequorea victoria.  
XX  
XX WO2004058973-A1.  
XX  
XX 15-JUL-2004.  
XX  
XX

PF 26-NOV-2003; 2003WO-RU000525.  
XX  
XX 26-DEC-2002; 2002US-0436857P.  
PR 02-APR-2003; 2003US-0459679P.  
XX  
XX (EYRO-) EYROGEN STOCK CO.  
XX  
XX Barsova EV, Lukyanov SA;  
XX WPI; 2004-525887/50.  
XX  
XX New isolated nucleic acid molecule encoding a fluorescent protein, useful  
XX for labeling biomolecules, cell, or cell organelles, or for identifying  
XX expression of a gene in a biological specimen.  
XX  
XX Example 1; Fig 1; 55pp; English.  
XX  
XX The present invention relates to novel coding sequences (1) encoding  
XX green fluorescent proteins (II). The fluorescent protein coding sequences  
XX are isolated from an organism from phylum Arthropoda, from subclass  
XX Copepoda, or from family Pontellidae. The fluorescent protein coding  
XX sequences are useful for labelling biomolecules, cell, or cell  
XX organelles. They are also useful for identifying expression of a gene in  
XX a biological specimen or for generating transforming constructs including  
XX transgenic organisms or site-specific gene modifications in cell lines.  
XX The fluorescent proteins are also useful in high throughput screening  
XX assays. The present sequence was used in a sequence alignment with the  
XX green fluorescent proteins of the invention.  
XX  
XX Sequence 225 AA;  
SQ

Query Match 100.0%; Score 1214; DB 8; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSKKNVKEFMFKRMKMEGTNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60  
DB 1 MRSKKNVKEFMFKRMKMEGTNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60  
QY 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120  
QY 121 KVFIFGVNFPSPDGPVWQKKTMGWEASTERLYPRDGYLKGSIHKALKKDGCHYLVEFKSI 180  
DB 121 KVFIFGVNFPSPDGPVWQKKTMGWEASTERLYPRDGYLKGSIHKALKKDGCHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 19  
ADX26534 standard; protein; 225 AA.  
ID ADX26534  
XX  
XX ADX26534;  
AC  
XX  
XX 21-APR-2005 (first entry)  
DT  
XX  
XX Diacosoma DsRed RFP protein, seq id 12.  
DE  
XX  
XX Phosphorylation; detection; red fluorescent protein; RFP.  
XX  
XX Diacosoma sp.  
XX  
XX US2005026234-A1.  
XX  
XX 03-FEB-2005.  
XX  
XX 26-MAY-2004; 2004US-00857622.  
XX  
XX 31-JAN-1996; 96US-00594575.  
XX  
XX

PR 31-JAN-1997; 97US-00792553.  
PR 13-SEP-1999; 99US-00396003.  
PR 24-MAY-2001; 2001US-00865291.  
XX (VIOL/) VIOLIN J D.  
XX (NEWTON/) NEWTON A C.  
XX (TSIE/) TSIE R Y.  
XX (ZHANG/) ZHANG J.  
FI Violin JD, Newton AC, Tsien RY, Zhang J;  
XX WPI; 2005-141377/15.  
XX N-PSDB; ADX26533.  
PT New chimeric phosphorylation indicator comprises a first fluorescent  
PT protein, phosphoinositide binding domain with PHA2 sequence and protein  
PT kinase C-phosphorylatable domain, useful for detecting protein kinase C  
PT or phosphates.  
PS Disclosure; SEQ ID NO 12; 68pp; English.  
XX The invention relates to a chimeric phosphorylation indicator (CPI) (I)  
XX comprising in operative linkage, a first fluorescent protein, a  
XX phosphoinositide binding domain with an PHA2 sequence defined in the  
XX specification, a protein kinase C (PKC)-phosphorylatable domain, and a  
XX second fluorescent protein. The first and second fluorescent proteins are  
XX chosen from green fluorescent proteins (GFPs) and red fluorescent  
XX proteins (RFPs). The first and second fluorescent proteins exhibit a  
XX detectable resonance energy transfer, when the first fluorescent protein  
XX is excited. The PKC-phosphorylatable domain and phosphoinositide binding  
XX domain do not substantially emit light to excite the second fluorescent  
XX protein. (I) is useful for detecting a protein kinase C (PKC) or  
XX phosphates in a sample (such as biological sample e.g. cell, tissue  
XX sample, or their extracts). (I) is useful for detecting a kinase  
XX inhibitor or phosphatase inhibitor. The method of the invention is  
XX adapted to high throughput analysis. The current sequence represents the  
XX amino acid sequence of Discosoma Dered RFP.  
SQ Sequence 225 AA;  
Query Match 100.0%; Score 1214; DB 9; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVIEKPRFKRMGTNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFAMDI 60  
DB 1 MRSSKNVIEKPRFKRMGTNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFAMDI 60  
QY 61 LSPQFQYGSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFTGVNFPSPDGPVWQKKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
DB 121 KYKFTGVNFPSPDGPVWQKKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
RESULT 20  
ABB08821  
ID ABB08821 standard; protein; 487 AA.  
XX ABB08821;  
XX  
XX 28-MAY-2002 (first entry)  
XX  
XX Autofluorescent fusion protein A SEQ ID NO 1.  
XX  
XX Autofluorescent; fusion protein; proteolytic activity; fluorophore;  
XX protease; assay; protease cleavage.

XX  
OS Synthetic.  
XX  
XX WO200212543-A2.  
XX  
XX 14-FEB-2002.  
XX  
XX 07-AUG-2001; 2001WO-EP009112.  
XX  
XX 07-AUG-2000; 2000DE-01038382.  
XX  
XX (DIRE-) DIREVO BIOTECH AG.  
XX  
XX Kuhlmann R, Koltermann A, Ketting U, Schwille P;  
XX WPI; 2002-269094/31.  
XX  
XX New autofluorescent fusion protein, useful for determining protease and  
XX protease-inhibiting activity, comprises two different proteins linked by  
XX protease cleavage site.  
XX  
XX Example; Page 11; 35pp; German.  
XX  
XX The invention relates to an autofluorescing fusion protein (I, ABB08821-  
XX ABB08823) comprising: (i) a first autofluorescing protein (Ia); (ii) a  
XX segment containing a protease cleavage site; and (iii) at least one  
XX different autofluorescing protein (Ib). Essentially no fluorescent energy  
XX transfer occurs between (Ia) and (Ib). (I) is used for detecting and  
XX quantifying protease (or protease-inhibitory) activity in liquid samples  
XX or cells, particularly for screening-based optimization (or generation)  
XX of biomolecules with proteolytic activity. (I) can be prepared in  
XX cellular or cell-free systems and makes possible intracellular analysis  
XX of protease activity. Preparation of (I) does not require regioselective  
XX coupling of fluorophores to polypeptides and any selected protease  
XX cleavage site can be incorporated  
XX  
SQ Sequence 487 AA;  
Query Match 100.0%; Score 1214; DB 5; Length 487;  
Best Local Similarity 100.0%; Pred. No. 3.2e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVIEKPRFKRMGTNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFAMDI 60  
DB 263 MRSSKNVIEKPRFKRMGTNGHFEIEGEGRPYEGHNTYVLTAKTKGGPLPFAMDI 322  
QY 61 LSPQFQYGSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
DB 323 LSPQFQYGSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 382  
QY 121 KYKFTGVNFPSPDGPVWQKKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
DB 383 KYKFTGVNFPSPDGPVWQKKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 442  
QY 443 YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTGRHHLFL 487  
DB 443 YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTGRHHLFL 487  
RESULT 21  
ABB08822  
ID ABB08822 standard; protein; 506 AA.  
XX ABB08822;  
XX  
XX 28-MAY-2002 (first entry)  
XX  
XX Autofluorescent fusion protein B SEQ ID NO 2.  
XX  
XX Autofluorescent; fusion protein; proteolytic activity; fluorophore;  
XX protease; assay; protease cleavage.  
XX  
XX Synthetic.

XX WO200212543-A2.  
XX 14-FEB-2002.  
XX 07-AUG-2001; 2001WO-EP009112.  
XX PF 07-AUG-2000; 2000DE-01038382.  
XX PR 07-AUG-2000; 2000DE-01038382.  
XX PA (DIRE-) DIREVO BIOTECH AG.  
XX PI Kuhlmann R, Koltermann A, Ketting U, Schwille P;  
XX WPI; 2002-269094/31.  
XX New autofluorescent fusion protein, useful for determining protease and  
XX protease-inhibiting activity, comprises two different proteins linked by  
XX protease cleavage site.  
XX  
XX Claim 6; Page 11; 35pp; German.  
XX The invention relates to an autofluorescing fusion protein (I, ABB08821-  
XX ABB08823) comprising: (i) a first autofluorescing protein (Ia); (ii) a  
XX segment containing a protease cleavage site; and (iii) at least one  
XX different autofluorescing protein (Ib). Essentially no fluorescent energy  
XX transfer occurs between (Ia) and (Ib). (I) is used for detecting and  
XX quantifying protease (or protease-inhibitory) activity in liquid samples  
XX or cells, particularly for screening-based optimisation (or generation)  
XX of biomolecules with proteolytic activity. (I) can be prepared in  
XX cellular or cell-free systems and makes possible intracellular analysis  
XX of protease activity. Preparation of (I) does not require regioselective  
XX coupling of fluorophores to polypeptides and any selected protease  
XX cleavage site can be incorporated

SQ Sequence 506 AA;

Query Match 100.0%; Score 1214; DB 5; Length 506;  
Best Local Similarity 100.0%; Pred. No. 3.4e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIEKPMFKVMRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTGGGLPRAWDI 60  
DB 282 MRSSKNVIEKPMFKVMRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTGGGLPRAWDI 341  
QY 61 LSPQFOYGSKVYVKHPADIIDYKLSFPEGFKMERVMNPFEDGGVVTVDSSIQDGCFTY 120  
DB 342 LSPQFOYGSKVYVKHPADIIDYKLSFPEGFKMERVMNPFEDGGVVTVDSSIQDGCFTY 401  
QY 121 KYVFIGVNFPSDGPVWQKKTGMGEASTERLYPRDGVLKGEIHKALKKDGSHLVFEKSI 180  
DB 402 KYVFIGVNFPSDGPVWQKKTGMGEASTERLYPRDGVLKGEIHKALKKDGSHLVFEKSI 461  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYITVEQYERTEGRHHLFL 225  
DB 462 YMAKKPVQLPGYYVYVDSKLDITSHNEDYITVEQYERTEGRHHLFL 506

RESULT 22

ABB08823 ID ABB08823 standard; protein; 547 AA.

XX ABB08823;

XX 28-MAY-2002 (first entry)

XX Autofluorescent fusion protein C SEQ ID NO 3.

XX Autofluorescent; fusion protein; proteolytic activity; fluorophore;  
XX protease; assay; protease cleavage.

XX Synthetic.

XX WO200212543-A2.

XX 14-FEB-2002.  
XX 07-AUG-2001; 2001WO-EP009112.  
XX PF 07-AUG-2000; 2000DE-01038382.  
XX PR 07-AUG-2000; 2000DE-01038382.  
XX PA (DIRE-) DIREVO BIOTECH AG.  
XX PI Kuhlmann R, Koltermann A, Ketting U, Schwille P;  
XX WPI; 2002-269094/31.  
XX New autofluorescent fusion protein, useful for determining protease and  
XX protease-inhibiting activity, comprises two different proteins linked by  
XX protease cleavage site.  
XX  
XX Claim 6; Page 11; 35pp; German.  
XX The invention relates to an autofluorescing fusion protein (I, ABB08821-  
XX ABB08823) comprising: (i) a first autofluorescing protein (Ia); (ii) a  
XX segment containing a protease cleavage site; and (iii) at least one  
XX different autofluorescing protein (Ib). Essentially no fluorescent energy  
XX transfer occurs between (Ia) and (Ib). (I) is used for detecting and  
XX quantifying protease (or protease-inhibitory) activity in liquid samples  
XX or cells, particularly for screening-based optimisation (or generation)  
XX of biomolecules with proteolytic activity. (I) can be prepared in  
XX cellular or cell-free systems and makes possible intracellular analysis  
XX of protease activity. Preparation of (I) does not require regioselective  
XX coupling of fluorophores to polypeptides and any selected protease  
XX cleavage site can be incorporated

SQ Sequence 547 AA;

Query Match 100.0%; Score 1214; DB 5; Length 547;  
Best Local Similarity 100.0%; Pred. No. 3.8e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIEKPMFKVMRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTGGGLPRAWDI 60  
DB 323 MRSSKNVIEKPMFKVMRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTGGGLPRAWDI 382  
QY 61 LSPQFOYGSKVYVKHPADIIDYKLSFPEGFKMERVMNPFEDGGVVTVDSSIQDGCFTY 120  
DB 383 LSPQFOYGSKVYVKHPADIIDYKLSFPEGFKMERVMNPFEDGGVVTVDSSIQDGCFTY 442  
QY 121 KYVFIGVNFPSDGPVWQKKTGMGEASTERLYPRDGVLKGEIHKALKKDGSHLVFEKSI 180  
DB 443 KYVFIGVNFPSDGPVWQKKTGMGEASTERLYPRDGVLKGEIHKALKKDGSHLVFEKSI 502  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYITVEQYERTEGRHHLFL 225  
DB 503 YMAKKPVQLPGYYVYVDSKLDITSHNEDYITVEQYERTEGRHHLFL 547

RESULT 23

AAE28920 ID AAE28920 standard; protein; 225 AA.

XX AAE28920;

XX 27-DEC-2002 (first entry)

XX Discosoma sp. drfp583 (NFP-6) mutant protein, E5 (S197T).

XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
XX fluorescence activated cell sorting application; fluorescent timer;  
XX biosensor; fluorescence resonance energy transfer application; FRET;  
XX colouring agent; recombinant DNA application; analyte detection assay;  
XX sunscreen; second messenger detector; drfp583 protein; NFP-6; mutant;  
XX mutcin.

XX Discosoma sp.

OS Synthetic. Location/Qualifiers  
FH Key Misc-difference 197 /note= "wild-type Ser substituted with Thr"  
FT  
XX MO200268459-A2.  
XX PN  
XX PD 06-SEP-2002.  
XX PF 20-FEB-2002; 2002WO-US005749.  
XX PR 21-FEB-2001; 2001US-0270983P.  
XX PR 04-DEC-2001; 2001US-00006922.  
XX PA (CLON-) CLONTECH LAB INC.  
XX PI Lukyanov S, Lukyanov K, Yanushkevich Y, Savitsky A, Fradkov A;  
XX WPI: 2002-691654/74.  
XX  
XX PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
XX PS Disclosure; Page; 80pp; English.  
XX  
XX CC The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as screens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drfp583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Discosoma sp. drfp583 (NFP-6) wild-type protein shown as SEQ  
CC ID NO:8 (AMB28833) in page 70-71 of the specification  
XX  
XX SQ Sequence 225 AA;  
SQ  
Query Match 99.8%; Score 1211; DB 5; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVTKLKTGSGPLPFAMDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVTKLKTGSGPLPFAMDI 60  
QY 61 LSPQOYSGSKYVVKRPADIPYKLSFPEGFKMERVMNFBGCVTVTVTODSSLQDGCFTY 120  
DB 61 LSPQOYSGSKYVVKRPADIPYKLSFPEGFKMERVMNFBGCVTVTVTODSSLQDGCFTY 120  
QY 121 KVKFIGNVPSDGPVWQKKTGMGEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGNVPSDGPVWQKKTGMGEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 121 KVKFIGNVPSDGPVWQKKTGMGEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGNVPSDGPVWQKKTGMGEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKRPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKRPVOLPGYVVYVDTKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 24  
ID ABM00937 standard; protein; 225 AA.  
XX AC ABM00937;  
XX DT 15-JAN-2004 (first entry)  
XX DE Discosoma sp. red fluorescent protein (RFP), S197T.

XX KW Fluorescent protein; resonance energy transfer; pH; detection;  
KW red fluorescent protein; RFP; mutant; mutein.  
XX OS Discosoma sp.  
OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Misc-difference 197 /note= "wild-type Ser is substituted with Thr"  
XX PN  
XX PD 11-SEP-2003.  
XX PF 26-FEB-2001; 2001US-00794308.  
XX PR 26-FEB-2001; 2001US-00794308.  
XX PR 26-FEB-2001; 2001US-00794308.  
XX PA (TSIE/) TSIE R Y.  
XX PA (ZACH/) ZACHARIAS D A.  
XX PA (BAIRD/) BAIRD G S.  
XX PI Tsien RY, Zacharias DA, Baird GS;  
XX WPI: 2003-802418/75.  
XX  
XX PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX  
XX PS Example 2; Page; 0pp; English.  
XX  
XX CC The invention relates to a non-oligomerising fluorescent protein  
XX containing a mutation that reduces or eliminates its ability to  
XX oligomerise. The fluorescent protein gives more reliable fluorescence  
XX resonance energy transfer results and are useful to detect molecule  
XX interaction, enzymes, or sample pH. These are also used to identify  
XX agents or conditions that regulate expression of control sequences. The  
XX present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
XX Note: This sequence is not shown in the specification, however this  
XX sequence is constructed based on Discosoma sp. wild-type RFP protein  
XX shown in page 30-31 (ABM0918)  
XX  
XX SQ Sequence 225 AA;  
SQ  
Query Match 99.8%; Score 1211; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVTKLKTGSGPLPFAMDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVTKLKTGSGPLPFAMDI 60  
QY 61 LSPQOYSGSKYVVKRPADIPYKLSFPEGFKMERVMNFBGCVTVTVTODSSLQDGCFTY 120  
DB 61 LSPQOYSGSKYVVKRPADIPYKLSFPEGFKMERVMNFBGCVTVTVTODSSLQDGCFTY 120  
QY 121 KVKFIGNVPSDGPVWQKKTGMGEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGNVPSDGPVWQKKTGMGEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 121 KVKFIGNVPSDGPVWQKKTGMGEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGNVPSDGPVWQKKTGMGEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKRPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKRPVOLPGYVVYVDTKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 25  
ID ABM00929 standard; protein; 225 AA.  
XX AC ABM00929;  
XX DT 15-JAN-2004 (first entry)  
XX DE Discosoma sp. red fluorescent protein (RFP), S197T.

XX 15-JAN-2004 (first entry)  
XX  
XX Discosoma sp. red fluorescent protein (RFP), K83R.  
DE  
XX Fluorescent protein; resonance energy transfer; pH; detection;  
XX red fluorescent protein; RFP; mutant; muteln.  
XX  
XX Discosoma sp.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 83 /note="Wild-type Lys is substituted with Arg"  
FT  
FT  
XX US2003170911-A1.  
XX  
XX 11-SEP-2003.  
XX  
XX 26-FEB-2001; 2001US-00794308.  
XX  
XX 26-FEB-2001; 2001US-00794308.  
XX  
XX (TSIE/) TSIE R Y.  
XX (ZACH/) ZACHARIAS D A.  
XX (BAIR/) BAIRD G S.  
XX  
XX Tsien RY, Zacharias DA, Baird GS;  
XX  
XX MPI; 2003-802418/75.  
XX  
XX Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
XX enzymes, or sample pH.  
XX  
XX Example 2; Page; 0pp; English.  
XX  
XX The invention relates to a non-oligomerizing fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerize. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
CC shown in page 30-31 (ABW00918)  
XX  
XX  
SQ Sequence 225 AA;  
Query Match 99.8%; Score 1211; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVYKEFMPFKYVMEGTVNGHFEIIEGEGRPRYEGHNTVKLKVTGKGLPFPAMD 60  
DB 1 MRSSKNVYKEFMPFKYVMEGTVNGHFEIIEGEGRPRYEGHNTVKLKVTGKGLPFPAMD 60  
QY 61 LSPQFOYGSKYVYVHPADIPDYKLSPEEGFKMERVNNFEDGGVVTYTDSSLDGCFIY 120  
DB 61 LSPQFOYGSKYVYVHPADIPDYKLSPEEGFKMERVNNFEDGGVVTYTDSSLDGCFIY 120  
QY 121 KVFFIGVNPSPDGVWQKKTMGWEASTERLYPRDGYLKGSIHKLKLDGSHYLVFKSI 180  
DB 121 KVFFIGVNPSPDGVWQKKTMGWEASTERLYPRDGYLKGSIHKLKLDGSHYLVFKSI 180  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEOYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEOYERTEGRHHLFL 225  
RESULT 26

ABW00938  
ID ABW00938 standard; protein; 225 AA.  
XX  
XX  
XX ABW00938;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX Discosoma sp. red fluorescent protein (RFP), K70R.  
DE  
XX Fluorescent protein; resonance energy transfer; pH; detection;  
XX red fluorescent protein; RFP; mutant; muteln.  
XX  
XX Discosoma sp.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 70 /note="Wild-type Lys is substituted with Arg"  
FT  
FT  
XX US2003170911-A1.  
XX  
XX 11-SEP-2003.  
XX  
XX 26-FEB-2001; 2001US-00794308.  
XX  
XX 26-FEB-2001; 2001US-00794308.  
XX  
XX (TSIE/) TSIE R Y.  
XX (ZACH/) ZACHARIAS D A.  
XX (BAIR/) BAIRD G S.  
XX  
XX Tsien RY, Zacharias DA, Baird GS;  
XX  
XX MPI; 2003-802418/75.  
XX  
XX Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
XX enzymes, or sample pH.  
XX  
XX Example 2; Page; 0pp; English.  
XX  
XX The invention relates to a non-oligomerizing fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerize. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
CC shown in page 30-31 (ABW00918)  
XX  
XX  
SQ Sequence 225 AA;  
Query Match 99.8%; Score 1211; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVYKEFMPFKYVMEGTVNGHFEIIEGEGRPRYEGHNTVKLKVTGKGLPFPAMD 60  
DB 1 MRSSKNVYKEFMPFKYVMEGTVNGHFEIIEGEGRPRYEGHNTVKLKVTGKGLPFPAMD 60  
QY 61 LSPQFOYGSKYVYVHPADIPDYKLSPEEGFKMERVNNFEDGGVVTYTDSSLDGCFIY 120  
DB 61 LSPQFOYGSKYVYVHPADIPDYKLSPEEGFKMERVNNFEDGGVVTYTDSSLDGCFIY 120  
QY 121 KVFFIGVNPSPDGVWQKKTMGWEASTERLYPRDGYLKGSIHKLKLDGSHYLVFKSI 180  
DB 121 KVFFIGVNPSPDGVWQKKTMGWEASTERLYPRDGYLKGSIHKLKLDGSHYLVFKSI 180  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEOYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEOYERTEGRHHLFL 225

Db 181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 27

ADH34498

ADH34498 standard; protein; 225 AA.

XX

XX

ADH34498;

XX

11-MAR-2004 (first entry)

XX

XX

Discosoma sp. DsRed mutant S148A.

XX

Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;

KM Chitarrarian; Anthozoa; labelling; colouring agents; pigment;

KM analyte detection assay; selectable marker; sunscreen; selective filter;

KM fluorescence resonance energy transfer; FRET; biosensor;

KM whole cell marker; second messenger detector; in vivo marker;

KM fluorescence activated cell sorting; fluorescent timer;

KM red fluorescent protein; DsRed; mutant; mutcin.

XX

XX

Synthetic.

OS Discosoma sp.

XX

Key Location/Qualifiers

XX Misc-difference 146

FT /note= "Ala replaces wild-type Ser. Corresponds to GFP

FT residue 148"

XX

XX

WO2003057833-A2.

XX

17-JUL-2003.

XX

23-DEC-2002; 2002WO-US041418.

XX

26-DEC-2001; 2001US-0343128P.

XX

(CLON-) CLONTECH LAB INC.

XX

XX

Bulina ME, Chudakov D, Lukyanov KA;

PI

XX

WPI; 2003-607998/57.

XX

Novel nucleic acid encoding interconverted mutant of chromo-or

PT fluorescent protein which are useful as biosensors, coloring agents.

XX

XX

Example 1; Page; 56pp; English.

XX

The invention relates to interconverted mutants of chromoproteins (CP) or

CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is

CC derived from a Chitarrarian species, preferably a non-bioluminescent

CC Chitarrarian species, and most preferably an Anthozoa species. The

CC invention is based on the finding that although green fluorescent protein

CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of

CC homology, there are certain positions (referred to as 148, 165, 167 and

CC 203; numbering corresponds to GFP) that are occupied by noticeably

CC different residues in the two types of proteins. Mutagenesis of the

CC residues in these key positions in, for example, a fluorescent protein,

CC to those found in a chromoprotein is therefore proposed to confer

CC chromoprotein activity on the fluorescent protein mutant, with

CC chromoproteins being able to be converted into fluorescent proteins in a

CC similar manner. The invention also relates to expression constructs,

CC vectors, host cells and host cell progeny comprising a nucleic acid of

CC the invention; the recombinant production of an interconverted

CC chromoprotein or fluorescent protein mutant; and antibodies specific for

CC interconverted mutant proteins of the invention. The interconverted

CC mutants are useful in any application that employs a chromoprotein or

CC fluorescent protein. Fluorescent protein mutants having chromoprotein

CC activity can be useful as colouring agents in, for example, food

CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins

CC with chromoprotein activity are also useful as labels in biological

CC analyte detection assays, as selectable markers in recombinant DNA

CC applications (e.g. the production of transgenic cells and organisms), and

CC are also useful as sunscreens and selective filters. Chromoprotein

CC mutants having fluorescent protein activity useful in fluorescence

CC resonance energy transfer (FRET) applications, as biosensors in

CC prokaryotic and eukaryotic cells, as markers of whole cells to detect

CC changes in multicellular reorganisation and migration, as second

CC messenger detectors, as in vivo markers in animals (e.g., transgenic

CC animals), in fluorescence activated cell sorting applications, in

CC protease cleavage assays, and in assays to determine the phospholipid

CC composition in biological membranes. Proteins with fluorescent protein

CC activity can also be used as fluorescent timers, where the switch of one

CC fluorescent colour to another (e.g., green to red) is concomitant with

CC the ageing of the protein and is useful for determination of the

CC activation or deactivation of gene expression. The present sequence

CC represents a Discosoma sp. red fluorescent protein DsRed mutant generated

CC in an example of the invention. The present sequence is not shown in the

CC specification, but was derived from the wild-type DsRed sequence

CC (ADH34489) shown in Fig 1 and the information provided on page 42.

XX

XX

SQ Sequence 225 AA;

XX

Query Match 99.8%; Score 1211; DB 7; Length 225;

Best Local Similarity 99.6%; Pred. No. 2.2e-127;

Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVTKEMFRFVYRMETVNGHEFEIEGEGRPYEGHNTVYLKTKTGAGPLFPADI 60

Db 1 MRSSKNVTKEMFRFVYRMETVNGHEFEIEGEGRPYEGHNTVYLKTKTGAGPLFPADI 60

QY 61 LSPQFQYSKYVVKHPADIPYKKLSPEGFKMERVWNPEDGVVTVTQDSSLQDGCFTY 120

Db 61 LSPQFQYSKYVVKHPADIPYKKLSPEGFKMERVWNPEDGVVTVTQDSSLQDGCFTY 120

QY 121 KYKFTGVNFPSPDGPVMQKKTGWEASTERLYPRDSVLGELTHKALKDGGHYLVFPSI 180

Db 121 KYKFTGVNFPSPDGPVMQKKTGWEASTERLYPRDSVLGELTHKALKDGGHYLVFPSI 180

QY 181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

Db 181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 28

ADH34499

ID ADH34499 standard; protein; 225 AA.

XX

XX

ADH34499;

XX

XX

11-MAR-2004 (first entry)

XX

XX

Discosoma sp. DsRed mutant S203A.

XX

Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;

KM Chitarrarian; Anthozoa; labelling; colouring agents; pigment;

KM analyte detection assay; selectable marker; suncreen; selective filter;

KM fluorescence resonance energy transfer; FRET; biosensor;

KM whole cell marker; second messenger detector; in vivo marker;

KM fluorescence activated cell sorting; fluorescent timer;

KM red fluorescent protein; DsRed; mutant; mutcin.

XX

XX

Synthetic.

OS Discosoma sp.

XX

Key Location/Qualifiers

XX Misc-difference 197

FT /note= "Ala replaces wild-type Ser. Corresponds to GFP

FT residue 203"

XX

XX

WO2003057833-A2.

XX

17-JUL-2003.

XX

23-DEC-2002; 2002WO-US041418.

FR 26-DEC-2001; 2001US-0343128P.  
 XX (CLON-) CLONTECH LAB INC.  
 PA  
 PI Bulina ME, Chudakov D, Lukyanov KA;  
 DR WPI; 2003-607998/57.  
 XX Novel nucleic acid encoding interconverted mutant of chromo-or  
 PT fluorescent protein which are useful as biosensors, coloring agents.  
 XX  
 PS Example 1, Page; 56pp; English.  
 CC The invention relates to interconverted mutants of chromoproteins (CP) or  
 CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
 CC derived from a Cnidarian species, preferably a non-bioluminescent  
 CC Cnidarian species, and most preferably an Anthozoan species. The  
 CC invention is based on the finding that although green fluorescent protein  
 CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
 CC homology, there are certain positions (referred to as 148, 165, 167 and  
 CC 203; numbering corresponds to GFP) that are occupied by noticeably  
 CC different residues in the two types of proteins. Mutagenesis of the  
 CC residues in these key positions in, for example, a fluorescent protein,  
 CC to those found in a chromoprotein is therefore proposed to confer  
 CC chromoprotein activity on the fluorescent protein mutant, with  
 CC chromoproteins being able to be converted into fluorescent proteins in a  
 CC similar manner. The invention also relates to expression constructs,  
 CC vectors, host cells and host cell progeny comprising a nucleic acid of  
 CC the invention; the recombinant production of an interconverted  
 CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
 CC interconverted mutant proteins of the invention. The interconverted  
 CC mutants are useful in any application that employs a chromoprotein or  
 CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
 CC activity can be useful as colouring agents in, for example, food  
 CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
 CC with chromoprotein activity are also useful as labels in biological  
 CC analyte detection assays, as selectable markers in recombinant DNA  
 CC applications (e.g. the production of transgenic cells and organisms), and  
 CC are also useful as sunscreens and selective filters. Chromoprotein  
 CC mutants having fluorescent protein activity useful in fluorescence  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
 CC changes in multicellular reorganisation and migration, as second  
 CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
 CC animals), in fluorescence activated cell sorting applications, in  
 CC protease cleavage assays, and in assays to determine the phospholipid  
 CC composition in biological membranes. Proteins with fluorescent protein  
 CC activity can also be used as fluorescent timers, where the switch of one  
 CC fluorescent colour to another (e.g., green to red) is concomitant with  
 CC the ageing of the protein and is useful for determination of the  
 CC activation or deactivation of gene expression. The present sequence  
 CC represents a Discosoma sp. red fluorescent protein Dered mutant generated  
 CC in an example of the invention. The present sequence is not shown in the  
 CC specification, but was derived from the wild-type Dered sequence  
 CC (ADH34469) shown in Fig 1 and the information provided on page 42.  
 CC  
 XX Sequence 225 AA:  
 SO  
 Query Match 99.8%; Score 1211; DB 7; Length 225;  
 Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
 Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRSSKNVKEFMKPKVMGEGTVNGHFEIIEGEGRGPRYEAGHTVTKLVTKGGLPFPAMD1 60  
 DB 1 MRSSKNVKEFMKPKVMGEGTVNGHFEIIEGEGRGPRYEAGHTVTKLVTKGGLPFPAMD1 60  
 QY 61 LSPQFOYGSKYVYVGPADIPDYKKLSPPSGPKWERNVNFEDGGVVTYVTDSSLDGCFIY 120  
 DB 61 LSPQFOYGSKYVYVGPADIPDYKKLSPPSGPKWERNVNFEDGGVVTYVTDSSLDGCFIY 120  
 QY 121 KYKPIGVNPFSDGPRVWQKKTVMGWASATERYLXPRDGVTKGSIHVKLKLDGGHVLVERKSI 180  
 DB 121 KYKPIGVNPFSDGPRVWQKKTVMGWASATERYLXPRDGVTKGSIHVKLKLDGGHVLVERKSI 180

QY 181 YMAKPPVQLPGYVYVDSKLDITSHNEDTYIVEQYERTEGRHHLPL 225  
 DB 181 YMAKPPVQLPGYVYVDSKLDITSHNEDTYIVEQYERTEGRHHLPL 225  
 RESULT 29  
 AAE28919 standard; protein, 225 AA.  
 ID AAE28919;  
 AC AAE28919;  
 XX 27-DEC-2002 (first entry)  
 DE Discosoma sp. drPP583 (NFP-6) mutant protein, ES (V105A).  
 KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
 KW fluorescence activated cell sorting application; fluorescent timer;  
 KW biosensor; fluorescence resonance energy transfer application; FRET;  
 KW colouring agent; recombinant DNA application; analyte detection assay;  
 KW sunscren; second messenger detector; drPP583 protein, NFP-6; mutant;  
 KW mutain.  
 XX Discosoma sp.  
 OS Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT Misc-difference 105 /note="Wild-type Val substituted with Ala"  
 FT  
 XX WO20026459-A2.  
 XX 06-SEP-2002.  
 XX 20-FEB-2002; 2002WO-US005749.  
 XX 21-FEB-2001; 2001US-0270983P.  
 XX 04-DEC-2001; 2001US-00006922.  
 XX (CLON-) CLONTECH LAB INC.  
 PA Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
 PI WPI; 2002-691654/74.  
 DR  
 XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
 PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
 PT analyte detection assays or fluorescence activated cell sorting  
 PT applications.  
 XX  
 PS Disclosure; Page; 80pp; English.  
 XX The invention relates to nucleic acid molecules encoding non-aggregating  
 CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
 CC useful in analyte detection assays, as colouring agents, as markers in  
 CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, in screening assays, as second  
 CC messenger detectors, in fluorescence activated cell sorting applications,  
 CC in protease cleavage assays or as fluorescent timers. The present  
 CC sequence is Discosoma sp. drPP583 (NFP-6) mutant protein of the  
 CC invention. Note: This sequence is not shown in the specification, but is  
 CC derived from Discosoma sp. drPP583 (NFP-6) wild-type protein shown as SEQ  
 CC ID NO:8 (AAE28833) in page 70-71 of the specification  
 CC  
 XX Sequence 225 AA:  
 SO  
 Query Match 99.7%; Score 1210; DB 5; Length 225;  
 Best Local Similarity 99.6%; Pred. No. 2.9e-127;  
 Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MRSSKNVKEFMKPKVMGEGTVNGHFEIIEGEGRGPRYEAGHTVTKLVTKGGLPFPAMD1 60  
 DB 1 MRSSKNVKEFMKPKVMGEGTVNGHFEIIEGEGRGPRYEAGHTVTKLVTKGGLPFPAMD1 60

DB 1 MRSSKNVKEFMRFKVRNEGTVNGHEFEIEBEGGRPYEGHNTVYKLTQTKGGPLPFANDI 60  
QY 61 LSPQOYGSKYVYVHPADI PDYKLSFPEGFKWERNVNFEDGGVTVTVQDSSLQDGCFTY 120  
DB 61 LSPQOYGSKYVYVHPADI PDYKLSFPEGFKWERNVNFEDGGVTVTVQDSSLQDGCFTY 120  
QY 121 KKFVIGVNFPSDGPVMOCKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFNSI 180  
DB 121 KKFVIGVNFPSDGPVMOCKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFNSI 180  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
RESULT 30  
ID AWM00930 standard; protein; 225 AA.  
AC AWM00930;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein (RFP), K83E.  
XX  
KW Fluorescent protein; resonance energy transfer; pH; detection;  
KM red fluorescent protein; RFP; mutant; mutain.  
XX  
OS Discosoma sp.  
OS Synthetic.  
XX  
FH Key location/Qualifiers  
FT Misc-difference 83 /note= "Wild-type Lys is substituted with Glu"  
XX  
PN US2003170911-A1.  
XX  
PD 11-SEP-2003.  
XX  
PF 26-FEB-2001; 2001US-00794308.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
XX  
PA (TSIE/) TSIE R Y.  
XX (ZACH/) ZACHARIAS D A.  
PA (BAIR/) BAIRD G S.  
XX  
PI Tsien RY, Zacharias DA, Baird GS,  
XX  
DR WPI, 2003-802418/75.  
XX  
PT Fluorescent proteins containing a mutation that reduces or eliminates its  
XX ability to oligomerize which gives more reliable fluorescence resonance  
XX energy transfer results and are useful to detect molecule interaction,  
XX enzymes, or sample pH.  
XX  
PS Example 2; Page; 0pp; English.  
XX  
CC The invention relates to a non-oligomerizing fluorescent protein  
XX containing a mutation that reduces or eliminates its ability to  
XX oligomerize. The fluorescent protein gives more reliable fluorescence  
XX resonance energy transfer results and are useful to detect molecule  
XX interaction, enzymes, or sample pH. These are also used to identify  
XX agents or conditions that regulate expression of control sequences. The  
XX present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
XX shown in page 30-31 (AWM00918)  
XX  
SQ Sequence 225 AA;  
Query Match 99.7%; Score 1210; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.9e-127;

Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRNEGTVNGHEFEIEBEGGRPYEGHNTVYKLTQTKGGPLPFANDI 60  
DB 1 MRSSKNVKEFMRFKVRNEGTVNGHEFEIEBEGGRPYEGHNTVYKLTQTKGGPLPFANDI 60  
QY 61 LSPQOYGSKYVYVHPADI PDYKLSFPEGFKWERNVNFEDGGVTVTVQDSSLQDGCFTY 120  
DB 61 LSPQOYGSKYVYVHPADI PDYKLSFPEGFKWERNVNFEDGGVTVTVQDSSLQDGCFTY 120  
QY 121 KKFVIGVNFPSDGPVMOCKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFNSI 180  
DB 121 KKFVIGVNFPSDGPVMOCKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFNSI 180  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
RESULT 31  
ID AAG65510 standard; protein; 226 AA.  
AC AAG65510;  
XX  
DT 30-NOV-2001 (first entry)  
XX  
DE Anthozoan red fluorescent protein synthetic mutant sequence.  
XX  
KW Anthozoan red fluorescent protein synthetic marker; FRET; mutant.  
XX  
OS Synthetic.  
XX  
PN WO200162919-A1.  
XX  
PD 30-AUG-2001.  
XX  
PF 13-FEB-2001; 2001WO-US004625.  
XX  
PR 23-FEB-2000; 2000US-0184732P.  
XX  
PA (AURO-) AURORA BIOSCIENCES CORP.  
XX  
PI Nelson D, Zamatra E, Tsien R,  
XX  
DR WPI, 2001-557704/62.  
XX  
DR N-PSDB; AAH47656.  
XX  
PT Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise  
XX functional red fluorescent proteins, and the encoding nucleic acids, with  
XX key mutations for improving the proteins function.  
XX  
PS Disclosure; Page 87; 90pp; English.  
XX  
CC The invention provides a nucleic acid encoding functional red fluorescent  
XX CC protein (II) that differs from the sequence of an Anthozoan red  
XX fluorescent protein by at least one amino acid substitution, and with  
XX different fluorescent properties. The red fluorescent protein of the  
XX invention can be expressed by standard recombinant methodology. (II) are  
XX used a fluorescent markers and FRET partners. It is used for identifying  
XX protein-protein interactions. (II) is also suitable for multiplexed  
XX fluorescent analysis and FRET-based applications using existing Aequorea  
XX fluorescent proteins. (II) has improved brightness, reduced spectral  
XX cross talk, and is rapidly and efficiently expressed in mammalian cells.  
XX The key mutations in the encoding nucleic acids provide improved folding,  
XX brightness, and create (II) with sharper, more defined excitation and  
XX emission peaks when expressed in mammalian cells. The present sequence  
XX represents the amino acid sequence of an improved synthetic mutant of an  
XX anthozoan fluorescent protein  
XX  
SQ Sequence 226 AA;  
Query Match 99.7%; Score 1210; DB 4; Length 226;



Beet Local Similarity 99.6%; Pred. No. 2.9e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVMEGTNGHEFEIEGEGRPYEGHNTVTKLKYTKGGPLPFANDI 60  
:|||||  
DB 2 VRSSKNVKEFMRFKVMEGTNGHEFEIEGEGRPYEGHNTVTKLKYTKGGPLPFANDI 61

QY 61 LSPQOYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
:|||||  
DB 62 LSPQOYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 121

QY 121 KYKFIGVNFPSDGPVMQKKTWMEASTERYLPDGLVKGELHKALKLKDGGHYLVEFKSI 180  
:|||||  
DB 122 KYKFIGVNFPSDGPVMQKKTWMEASTERYLPDGLVKGELHKALKLKDGGHYLVEFKSI 181

QY 181 YMAKKEPVQLPGYYVDSKLDITSHNEDYTYVEQYRTGRHHLFL 225  
:|||||  
DB 182 YMAKKEPVQLPGYYVDSKLDITSHNEDYTYVEQYRTGRHHLFL 226

RESULT 32  
ADE24109  
ID ADE24109 standard; protein; 242 AA.  
XX ADE24109;  
AC  
XX  
XX  
DT 29-JAN-2004 (first entry)  
XX  
XX  
DE Discosoma dp Dered1 reporter protein.  
XX  
XX site-specific DNA recombination; Cre recombinase;  
KW protein-transduction domain; gene activation; gene inactivation;  
KM chromosomal translocation; reporter protein.  
XX  
XX Discosoma dp.  
OS  
XX  
XX WO2003070931-A2.  
PN  
PD 28-AUG-2003.  
XX  
XX 19-FEB-2003; 2003WO-EP001680.  
PF  
XX 21-FEB-2002; 2002DE-01007313.  
PR 16-JUL-2002; 2002DE-01032196.  
XX  
XX (VIST-) VISION 7 GMBH.  
PA  
PI Baum C, Will E, Osterlag W, Klump H, Schiedlmeier B;  
XX  
XX WPI; 2003-767353/72.  
XX  
XX In vitro or in vivo site-specific DNA recombination, useful e.g. for gene  
PT inactivation, using Cre recombinase that lacks heterologous protein-  
PT transduction domain.  
XX  
XX Claim 22; SEQ ID NO 6; 84bp; German.  
PS  
XX

This invention describes a novel method of site-specific DNA recombination in eukaryotic cells in vitro comprising using a bacteriophage P1 Cre recombinase that lacks heterologous protein-transduction domains. The invention also contains a reporter system for detecting site-specific DNA recombination in eukaryotic cells. Recombination with Cre recombinase lacking heterologous protein-transduction domains is used for therapeutic site-specific recombination in eukaryotic cells, in vivo or in vitro, e.g. for partial inactivation or activation of genes, also to create a chromosomal translocation. The modified recombinase can enter intact (not electrically or chemically treated) cells when added to culture medium and catalyzes recombination events in the nucleus with recombination rate over 50%, after a single application. The Cre-recombinase does not need to be supplied from nucleic acid introduced into the cell, so its effect is immediate but short-lived, limiting cytotoxic effects, eliminating the risk of non-specific integration of DNA into the genome and providing quick results.

CC This sequence represents the Discosoma sp reporter protein Dered1 used in  
CC the method described in the invention.  
XX  
XX SQ Sequence 242 AA;  
XX  
XX Query Match 99.7%; Score 1210; DB 7; Length 242;  
Beet Local Similarity 99.6%; Pred. No. 3.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVMEGTNGHEFEIEGEGRPYEGHNTVTKLKYTKGGPLPFANDI 60  
:|||||  
DB 15 VRSSKNVKEFMRFKVMEGTNGHEFEIEGEGRPYEGHNTVTKLKYTKGGPLPFANDI 74

QY 61 LSPQOYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
:|||||  
DB 75 LSPQOYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 134

QY 121 KYKFIGVNFPSDGPVMQKKTWMEASTERYLPDGLVKGELHKALKLKDGGHYLVEFKSI 180  
:|||||  
DB 135 KYKFIGVNFPSDGPVMQKKTWMEASTERYLPDGLVKGELHKALKLKDGGHYLVEFKSI 194

QY 181 YMAKKEPVQLPGYYVDSKLDITSHNEDYTYVEQYRTGRHHLFL 225  
:|||||  
DB 195 YMAKKEPVQLPGYYVDSKLDITSHNEDYTYVEQYRTGRHHLFL 239

RESULT 33  
ADL18132  
ID ADL18132 standard; protein; 545 AA.  
XX  
XX ADL18132;  
AC  
XX  
XX  
DT 06-MAY-2004 (first entry)  
XX  
XX  
DE RFP:PS (N1a protease):AtOEP7.GFP fusion protein SEQ ID NO:52.  
XX  
XX  
XX chimeric protein; signal protein; trafficking signal targeting;  
KW proteolytic cleavage site; protease; protease inhibitor.  
XX  
XX Arabidopsis thaliana.  
OS  
XX Tobacco vein motting virus.  
OS  
XX Synthetic.  
OS  
XX WO2003014381-A1.  
PN  
XX 20-FEB-2003.  
PD  
XX 08-AUG-2002; 2002WO-KR001515.  
PF  
XX 10-AUG-2001; 2001KR-00048123.  
PR  
XX (AHRA-) AHRAM BIOSYSTEMS INC.  
PA  
XX  
XX Hwang I, Kim DH, Lee YJ;  
XX  
XX WPI; 2003-256596/25.  
DR N-PSDB; ADL18131.  
XX  
XX  
XX New chimeric protein, useful for detecting protease inhibitors inside the  
PT cell or tissue.  
PT  
XX  
XX Example 2; SEQ ID NO 52; 214bp; English.  
PS  
XX

The present invention describes a chimeric protein comprising at least one signal protein that has a trafficking signal targeting to a subcellular organelle and at least one proteolytic cleavage site for a protease. The chimeric protein is constructed, so that: (a) the trafficking signals of all the signal proteins are inactivated by linking the proteolytic site or a signal masking protein through the proteolytic site to the N- or C- terminus of the signal proteins, and so the chimeric protein is present in cytosol; (b) the trafficking signal of at least one signal protein is activated when the proteolytic cleavage site is cleaved by the protease, and as a result at least one fragment protein that

includes the activated signal protein is a transported to a subcellular  
 organelle; and (c) the chimeric protein is labelled with at least one  
 fluorescent protein and the position and intensity distribution of the  
 fluorescent label signal in the cell is altered depending on the cleavage  
 by the protease. Also described: (1) a recombinant gene comprising a  
 nucleic acid sequence encoding the chimeric protein which is constructed  
 to express the chimeric protein in a cell; (2) a cell transformed with  
 the recombinant gene or vector; (3) analysing the activity of a protease  
 in vivo; (4) screening protease inhibitors in vivo; (5) a system for  
 detecting a protease inside a cell; (6) a nucleic acid comprising the  
 sequence encoding the chimeric protein for detecting protease activity in  
 a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting  
 a protease inside a cell comprising the chimeric protein or the vector;  
 (9) detecting a protease inside a cell or tissue; and (10) detecting a  
 protease inhibitor in vivo. The chimeric protein is useful for detecting  
 protease inhibitors inside the cell or tissue. The present sequence  
 represents a fusion protein, which is used in the exemplification of the  
 present invention.

SQ Sequence 545 AA;

Query Match 99.7%; Score 1210; DB 7; Length 545;  
 Best Local Similarity 99.6%; Pred. No. 1.1e-126;  
 Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 MRSSKNVIKEFRPKVMEGTVNGHEFEIEBGEGRPYEGHNTVYKLTQYKGGPLPFAMD 60  
 :|||||  
 2 VRSSKNVIKEFRPKVMEGTVNGHEFEIEBGEGRPYEGHNTVYKLTQYKGGPLPFAMD 61  
 DB 61 LSPQFQYGSKYVYVHPADIPDYKLSPEEGFKWERNVNFEDGGVTVYQDSSLQDGCFTY 120  
 QY 62 LSPQFQYGSKYVYVHPADIPDYKLSPEEGFKWERNVNFEDGGVTVYQDSSLQDGCFTY 121  
 DB 121 KYKFTGVNFPSPDGVPMQKTMGWEASTERYLPDGVLKGEIHKALKDGGHYLVEFKSI 180  
 DB 122 KYKFTGVNFPSPDGVPMQKTMGWEASTERYLPDGVLKGEIHKALKDGGHYLVEFKSI 181  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHFL 225  
 QY 182 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHFL 226  
 DB

RESULT 34  
 ADL18156  
 ID ADL18156 standard; protein; 548 AA.

AC ADL18156;  
 DT 06-MAY-2004 (first entry)  
 DE RFP:PS(HIV-1 protease):AtoEP7:GFP fusion protein SEQ ID NO:76.  
 XX chimeric protein; signal protein; trafficking signal targeting;  
 KM proteolytic cleavage site; protease; protease inhibitor.  
 XX Arabidopsis thaliana.  
 OS Human immunodeficiency virus 1.  
 OS Synthetic.  
 XX WO2003014381-A1.  
 PN 20-FEB-2003.  
 PD 08-AUG-2002; 2002MO-KR001515.  
 PF 10-AUG-2001; 2001KR-00048123.  
 PR (AHRRA-) AHRAM BIOSYSTEMS INC.  
 PA Hwang I, Kim DH, Lee YJ;  
 XX WPI, 2003-256596/25.  
 DR N-PSDB; ADL18155.  
 DR

XX New chimeric protein, useful for detecting protease inhibitors inside the  
 PT cell or tissue.  
 PR

XX Example 3; SEQ ID NO 76; 214dp; English.

The present invention describes a chimeric protein comprising at least  
 one signal protein that has a trafficking signal targeting to a  
 subcellular organelle and at least one proteolytic cleavage site for a  
 protease. The chimeric protein is constructed, so that: (a) the  
 trafficking signals of all the signal proteins are inactivated by linking  
 the proteolytic site or a signal masking protein through the proteolytic  
 site to the N- or C- termini of the signal proteins; and so the chimeric  
 protein is present in cytosol; (b) the trafficking signal of at least one  
 signal protein is activated when the proteolytic cleavage site is cleaved  
 by the protease, and as a result at least one fragment protein that  
 includes the activated signal protein is transported to a subcellular  
 organelle; and (c) the chimeric protein is labelled with at least one  
 fluorescent protein and the position and intensity distribution of the  
 fluorescent label signal in the cell is altered depending on the cleavage  
 by the protease. Also described: (1) a recombinant gene comprising a  
 nucleic acid sequence encoding the chimeric protein which is constructed  
 to express the chimeric protein in a cell; (2) a cell transformed with  
 the recombinant gene or vector; (3) analysing the activity of a protease  
 in vivo; (4) screening protease inhibitors in vivo; (5) a system for  
 detecting a protease inside a cell; (6) a nucleic acid comprising the  
 sequence encoding the chimeric protein for detecting protease activity in  
 a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting  
 a protease inside a cell comprising the chimeric protein or the vector;  
 (9) detecting a protease inside a cell or tissue; and (10) detecting a  
 protease inhibitor in vivo. The chimeric protein is useful for detecting  
 protease inhibitors inside the cell or tissue. The present sequence  
 represents a fusion protein, which is used in the exemplification of the  
 present invention.

SQ Sequence 548 AA;

Query Match 99.7%; Score 1210; DB 7; Length 548;  
 Best Local Similarity 99.6%; Pred. No. 1.1e-126;  
 Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 MRSSKNVIKEFRPKVMEGTVNGHEFEIEBGEGRPYEGHNTVYKLTQYKGGPLPFAMD 60  
 :|||||  
 2 VRSSKNVIKEFRPKVMEGTVNGHEFEIEBGEGRPYEGHNTVYKLTQYKGGPLPFAMD 61  
 QY 61 LSPQFQYGSKYVYVHPADIPDYKLSPEEGFKWERNVNFEDGGVTVYQDSSLQDGCFTY 120  
 DB 62 LSPQFQYGSKYVYVHPADIPDYKLSPEEGFKWERNVNFEDGGVTVYQDSSLQDGCFTY 121  
 QY 121 KYKFTGVNFPSPDGVPMQKTMGWEASTERYLPDGVLKGEIHKALKDGGHYLVEFKSI 180  
 DB 122 KYKFTGVNFPSPDGVPMQKTMGWEASTERYLPDGVLKGEIHKALKDGGHYLVEFKSI 181  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHFL 225  
 DB 182 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHFL 226  
 DB

RESULT 35  
 AAE28922  
 ID AAE28922 standard; protein; 225 AA.

AC AAE28922;  
 DT 27-DEC-2002 (first entry)  
 DE Discosoma sp. drFP583 (NFP-6) mutant protein, E8 (N42H).  
 XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
 KM fluorescence activated cell sorting application; fluorescent timer;  
 KM biosensor; fluorescence resonance energy transfer application; FRRET;  
 KM colouring agent; recombinant DNA application; analyte detection assay;  
 KM sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
 KM

KW mutein.  
 XX  
 OS Discosoma sp.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 42  
 FT /note= "Wild-type Asn substituted with His"  
 XX  
 XX  
 PN MO20026459-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 20-FEB-2002; 2002WO-US005749.  
 XX  
 PR 21-FEB-2001; 2001US-0270983P.  
 PR 04-DEC-2001; 2001US-00006922.  
 XX  
 XX (CLON-) CLONTECH LAB INC.  
 XX  
 PI Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Fradkov A;  
 XX  
 DR WPI; 2002-691654/74.  
 XX  
 XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
 PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
 PT analyte detection assays or fluorescence activated cell sorting  
 PT applications.  
 XX  
 XX Disclosure; Page; 80pp; English.  
 XX  
 XX The invention relates to nucleic acid molecules encoding non-aggregating  
 CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
 CC useful in analyte detection assays, as colouring agents, as markers in  
 CC recombinant DNA transfer (FRET) applications, as biosensors in  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, in screening assays, as second  
 CC messenger detectors, in fluorescence activated cell sorting applications,  
 CC in protease cleavage assays or as fluorescent timers. The present  
 CC sequence is Discosoma sp. drp583 (NFP-6) mutant protein of the  
 CC invention. Note: This sequence is not shown in the specification, but is  
 CC derived from Discosoma sp. drp583 (NFP-6) wild-type protein shown as SEQ  
 CC ID NO:8 (AAE28833) in page 70-71 of the specification  
 XX  
 XX Sequence 225 AA:  
 SQ

Query Match 99.6%; Score 1209; DB 5; Length 225;  
 Best Local Similarity 99.6%; Pred. No. 3.8e-127;  
 Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKVVIVKEMFKRMKMEGTUNGHEPEIEGEGRPRYEGHNTVYKLVTKGGLPFAWDI 60  
 DB 1 MRSSKVVIVKEMFKRMKMEGTUNGHEPEIEGEGRPRYEGHNTVYKLVTKGGLPFAWDI 60  
 QY 61 LSPQFOYGSKVYVYKHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFFIY 120  
 DB 61 LSPQFOYGSKVYVYKHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFFIY 120  
 QY 121 KVFIFGNFSPSDGPNVQKKTWGEASTERLYPRDGVYKGEIHAKLKLKDGGHYLVFKSI 180  
 DB 121 KVFIFGNFSPSDGPNVQKKTWGEASTERLYPRDGVYKGEIHAKLKLKDGGHYLVFKSI 180  
 QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYITVEOYERTEGRHHFL 225  
 DB 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYITVEOYERTEGRHHFL 225

RESULT 36  
 ABM00936  
 ID ABM00936 standard; protein: 225 AA.  
 XX  
 AC ABM00936;  
 XX

DT 15-JAN-2004 (first entry)  
 XX  
 DB Discosoma sp. red fluorescent protein (RFP), Y120H.  
 XX  
 XX Fluorescent protein; resonance energy transfer; pH; detection;  
 KW red fluorescent protein; RFP; mutant; mutein.  
 XX  
 OS Discosoma sp.  
 OS Synthetic.  
 XX  
 OS  
 FH Key Location/Qualifiers  
 FT Misc-difference 120  
 FT /note= "Wild-type Tyr is substituted with His"  
 XX  
 XX  
 PN US2003170911-A1.  
 XX  
 PD 11-SEP-2003.  
 XX  
 PF 26-FEB-2001; 2001US-00794308.  
 XX  
 PR 26-FEB-2001; 2001US-00794308.  
 XX  
 PR 26-FEB-2001; 2001US-00794308.  
 XX  
 PA (TSIE/) TSIE R Y.  
 PA (ZACH/) ZACHARIAS D A.  
 PA (BAIR/) BAIRD G S.  
 XX  
 PI Tsien RY, Zacharias DA, Baird GS,  
 XX  
 DR WPI; 2003-802418/75.  
 XX  
 XX Fluorescent proteins containing a mutation that reduces or eliminates its  
 PT ability to oligomerize which gives more reliable fluorescence resonance  
 PT energy transfer results and are useful to detect molecule interaction,  
 PT enzymes, or sample pH.  
 XX  
 XX Example 2; Page; 0pp; English.  
 XX  
 XX The invention relates to a non-oligomerizing fluorescent protein  
 CC containing a mutation that reduces or eliminates its ability to  
 CC oligomerize. The fluorescent protein gives more reliable fluorescence  
 CC resonance energy transfer results and are useful to detect molecule  
 CC interaction, enzymes, or sample pH. These are also used to identify  
 CC agents or conditions that regulate expression of control sequences. The  
 CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
 CC Note: This sequence is not shown in the specification, however this  
 CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
 CC shown in page 30-31 (ABM00918)  
 XX  
 XX Sequence 225 AA:  
 SQ

Query Match 99.6%; Score 1209; DB 7; Length 225;  
 Best Local Similarity 99.6%; Pred. No. 3.8e-127;  
 Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKVVIVKEMFKRMKMEGTUNGHEPEIEGEGRPRYEGHNTVYKLVTKGGLPFAWDI 60  
 DB 1 MRSSKVVIVKEMFKRMKMEGTUNGHEPEIEGEGRPRYEGHNTVYKLVTKGGLPFAWDI 60  
 QY 61 LSPQFOYGSKVYVYKHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFFIY 120  
 DB 61 LSPQFOYGSKVYVYKHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFFIY 120  
 QY 121 KVFIFGNFSPSDGPNVQKKTWGEASTERLYPRDGVYKGEIHAKLKLKDGGHYLVFKSI 180  
 DB 121 KVFIFGNFSPSDGPNVQKKTWGEASTERLYPRDGVYKGEIHAKLKLKDGGHYLVFKSI 180  
 QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYITVEOYERTEGRHHFL 225  
 DB 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYITVEOYERTEGRHHFL 225

RESULT 37  
 ABM00931

```

ID  ABM00931 standard; protein; 225 AA.
XX
XX  ABM00931;
AC
XX  15-JAN-2004 (first entry)
DT
XX  Discosoma sp. red fluorescent protein (RFP), K83N.
DE
XX  Fluorescent protein; resonance energy transfer; pH; detection;
KW  red fluorescent protein; RFP; mutant; mutein.
XX
XX  Discosoma sp.
OS  Synthetic.
XX
XX  Key location/Qualifiers
FH  Misc-difference 83 /note= "Wild-type Lys is substituted with Asn"
FT
XX  US2003170911-A1.
XX  11-SEP-2003.
XX  26-FEB-2001; 2001US-00794308.
XX  26-FEB-2001; 2001US-00794308.
XX  26-FEB-2001; 2001US-00794308.
XX  (TSTE/) TSTEEN R Y.
XX  (ZACH/) ZACHARIAS D A.
XX  (BAIR/) BAIRD G S.
XX  Telen RY, Zacharias DA, Baird GS;
PI  WPI; 2003-802418/75.
DR
XX  Fluorescent proteins containing a mutation that reduces or eliminates its
PT  ability to oligomerize which gives more reliable fluorescence resonance
PT  energy transfer results and are useful to detect molecule interaction,
XX  enzyme, or sample pH.
XX
XX  Example 2; Page; 0pp; English.
XX
XX  The invention relates to a non-oligomerizing fluorescent protein
CC  containing a mutation that reduces or eliminates its ability to
CC  oligomerize. The fluorescent protein gives more reliable fluorescence
CC  resonance energy transfer results and are useful to detect molecule
CC  interaction, enzymes, or sample pH. These are also used to identify
CC  agents or conditions that regulate expression of control sequences. The
CC  present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.
CC  Note: This sequence is not shown in the specification, however this
CC  sequence is constructed based on Discosoma sp. wild-type RFP protein
CC  shown in page 30-31 (ABM00918)
XX
XX  Sequence 225 AA;
SQ
Query Match 99.6%; Score 1209; DB 7; Length 225;
Best Local Similarity 99.6%; Pred. No. 3.8e-127;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSSKNVTKERFRRFYRMGEVTNGHGFRIEGBGSRPYEGHNTVTKLKYTKGGPLPPAMDIL 60
DB 1 MSSKNVTKERFRRFYRMGEVTNGHGFRIEGBGSRPYEGHNTVTKLKYTKGGPLPPAMDIL 60
QY 61 LSPQOYGSKVYVKKPADIPDYKKLSFPEGFKMERVMNPNEDGGVTVVTDSSLDQGCFTYK 120
DB 61 LSPQOYGSKVYVKKPADIPDYKKLSFPEGFKMERVMNPNEDGGVTVVTDSSLDQGCFTYK 120
QY 121 KVFPGVNPSPDGPVMOCKTMGMEASTERTLYPRDGLKGEIHKALKLKDGGHYLVFESKIY 180
DB 121 KVFPGVNPSPDGPVMOCKTMGMEASTERTLYPRDGLKGEIHKALKLKDGGHYLVFESKIY 180
QY 181 YMAKPPVOLPGYVVYDSSKLDITSHNEDYTIYEQYERTBGRHHLFL 225
DB 181 YMAKPPVOLPGYVVYDSSKLDITSHNEDYTIYEQYERTBGRHHLFL 225

```

```

RESULT 38
ABM08835
ID  ABM08835 standard; protein; 226 AA.
XX
XX  ABM08835;
AC
XX  29-MAY-2002 (first entry)
DT
XX  Yeast optimised RFP SEQ ID NO 20.
DE
XX
XX  Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;
KW  Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;
KW  Escherichia coli; green fluorescent protein; biotechnology.
XX
XX  Anthozoa.
OS
XX  DE20001395-U1.
XX  15-MAR-2001.
XX  27-JAN-2000; 2000DE-02001395.
XX  27-JAN-2000; 2000DE-02001395.
XX  27-JAN-2000; 2000DE-02001395.
XX  (GPCB-) GPC BIOTECH AG.
XX  WPI; 2002-228394/29.
XX  N-PSDB; ABA95906.
XX
XX  New DNA encoding red fluorescent protein, useful as marker in
PT  biotechnology, has sequence optimized for expression in eukaryotes,
PT  especially yeast or plants.
XX
XX  Disclosure; Page 15; 19pp; German.
XX
XX  The invention relates to DNA (1) containing either sequence ABA95905 or
CC  sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein
CC  (yRFP). (1) are used to express red fluorescent protein (RFP) in
CC  eukaryotes, especially yeast, especially Saccharomyces cerevisiae and
CC  plants, especially dicotyledonous plants including Nicotiana tabacum or
CC  Arabidopsis thaliana and also in prokaryotes, especially bacteria,
CC  especially Escherichia coli. RFP is useful in the same way as green
CC  fluorescent protein but is more generally applicable in modern
CC  biotechnology. (1) are optimised for expression in yeast and so generate
CC  RFP at higher levels with stronger fluorescence and thus lowers the
CC  detection limit and gives a better signal-to-noise ratio. The present
CC  sequence is that of the yeast optimised RFP
XX
XX  Sequence 226 AA;
SQ
Query Match 99.6%; Score 1209; DB 5; Length 226;
Best Local Similarity 100.0%; Pred. No. 3.8e-127;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RSSKNVTKERFRRFYRMGEVTNGHGFRIEGBGSRPYEGHNTVTKLKYTKGGPLPPAMDIL 61
DB 3 RSSKNVTKERFRRFYRMGEVTNGHGFRIEGBGSRPYEGHNTVTKLKYTKGGPLPPAMDIL 62
QY 62 SPQOYGSKVYVKKPADIPDYKKLSFPEGFKMERVMNPNEDGGVTVVTDSSLDQGCFTYK 121
DB 63 SPQOYGSKVYVKKPADIPDYKKLSFPEGFKMERVMNPNEDGGVTVVTDSSLDQGCFTYK 122
QY 122 VKFPGVNPSPDGPVMOCKTMGMEASTERTLYPRDGLKGEIHKALKLKDGGHYLVFESKIY 181
DB 123 VKFPGVNPSPDGPVMOCKTMGMEASTERTLYPRDGLKGEIHKALKLKDGGHYLVFESKIY 182
QY 182 YMAKPPVOLPGYVVYDSSKLDITSHNEDYTIYEQYERTBGRHHLFL 225
DB 183 YMAKPPVOLPGYVVYDSSKLDITSHNEDYTIYEQYERTBGRHHLFL 226

```

ID	ABM00932	standard; protein; 225 AA.
XX	ABM00932;	
AC	15-JAN-2004	(first entry)
XX		
DE	Discozoma sp.	red fluorescent protein (RFP), K83P.
XX		
KM	Fluorescent protein; resonance energy transfer; pH; detection;	
KM	red fluorescent protein; RFP; mutant; muten.	
XX		
OS	Discozoma sp.	
XX	Synthetic.	
XX		
XX	Key	Location/Qualifiers
FT	Misc-difference 83	/note= "wild-type Lys is substituted with Pro"
PN	US2003170911-A1.	
XX		
PD	11-SEP-2003.	
XX		
PF	26-FEB-2001; 2001US-00794308.	
XX		
PR	26-FEB-2001; 2001US-00794308.	
XX		
PA	(TSIE// TSIE N Y.	
XX	(ZACH// ZACHARIAS D A.	
XX	(BAIR// BAIRD G S.	
PI	Tsien RY, Zacharias DA, Baird GS;	
DR	WPI; 2003-802418/75.	
XX		
PT	Fluorescent proteins containing a mutation that reduces or eliminates its	
PT	ability to oligomerize which gives more reliable fluorescence resonance	
PT	energy transfer results and are useful to detect molecule interaction,	
XX	enzymes, or sample pH.	
PS	Example 2; Page; Opp; English.	
XX		
CC	The invention relates to a non-oligomerizing fluorescent protein	
CC	containing a mutation that reduces or eliminates its ability to	
CC	oligomerize. The fluorescent protein gives more reliable fluorescence	
CC	resonance energy transfer results and are useful to detect molecule	
CC	interaction, enzymes, or sample pH. These are also used to identify	
CC	agents or conditions that regulate expression of control sequences. The	
CC	present sequence is Discozoma sp. red fluorescent protein (RFP) mutant.	
CC	Note: This sequence is not shown in the specification, however this	
CC	sequence is constructed based on Discozoma sp. wild-type RFP protein	
CC	shown in page 30-31 (ABM00918)	
XX		
SQ	Sequence 225 AA;	
QY	Query Match	99.5%; Score 1208; DB 7; Length 225;
DB	Best Local Similarity	99.6%; Pred. No. 4.9e-127;
DB	Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 MRSSNNVKEKMRFRVMEGVNNGHEFEIEGBGEPRFEGHNIVLKYTKGGRPFANDI	60
DB	1 MRSSNNVKEKMRFRVMEGVNNGHEFEIEGBGEPRFEGHNIVLKYTKGGRPFANDI	60
QY	61 LSPQRYQSSKYVKKRPADIPDYKLSPEEGFKMEKVMNPFEDGGVTVYQDSSLQDGCFTY	120
DB	61 LSPQRYQSSKYVKKRPADIPDYKLSPEEGFKMEKVMNPFEDGGVTVYQDSSLQDGCFTY	120
QY	121 KYKPIGVNPPSDGPNVQKTKMGWASTERLYPRDGLVKGELHKAALKDGGHYLVFKSI	180
DB	121 KYKPIGVNPPSDGPNVQKTKMGWASTERLYPRDGLVKGELHKAALKDGGHYLVFKSI	180
QY	181 YMAKPPVQLPGYYVDSKLDITSNHEDTIYVQYRTGRNHLFL	225

Query Match	Best Local Similarity	99.5%	Score 1208	DB 7	Length 225
Matches 224	Conservative	0	Mismatches 1	Indels 0	Gaps 0
1	MRSSKNVKEEMRFKVRMEGTVNGHEPEIEBGBEGRPYEGHNTVYLKVTYKGGLPFPAMD	1	MRSSKNVKEEMRFKVRMEGTVNGHEPEIEBGBEGRPYEGHNTVYLKVTYKGGLPFPAMD	1	MRSSKNVKEEMRFKVRMEGTVNGHEPEIEBGBEGRPYEGHNTVYLKVTYKGGLPFPAMD
61	LSPOQYQSKYVYKHPADIPDYKLSFPEGGRMEVVMFEDGGVTVTVQDSSLQDGCFTY	61	LSPOQYQSKYVYKHPADIPDYKLSFPEGGRMEVVMFEDGGVTVTVQDSSLQDGCFTY	61	LSPOQYQSKYVYKHPADIPDYKLSFPEGGRMEVVMFEDGGVTVTVQDSSLQDGCFTY
61	LSPOQYQSKYVYKHPADIPDYKLSFPEGGRMEVVMFEDGGVTVTVQDSSLQDGCFTY	61	LSPOQYQSKYVYKHPADIPDYKLSFPEGGRMEVVMFEDGGVTVTVQDSSLQDGCFTY	61	LSPOQYQSKYVYKHPADIPDYKLSFPEGGRMEVVMFEDGGVTVTVQDSSLQDGCFTY
121	KVKFPGVFPSPDGMQKKTGMGEASTERLVPDGLVKGELIKALKLKDGGHYLVEFKSI	121	KVKFPGVFPSPDGMQKKTGMGEASTERLVPDGLVKGELIKALKLKDGGHYLVEFKSI	121	KVKFPGVFPSPDGMQKKTGMGEASTERLVPDGLVKGELIKALKLKDGGHYLVEFKSI

```

DB      121  |||||
QY      181  YMAKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
DB      181  YMAKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 41
ABM00939
ID      ABM00939 standard; protein; 225 AA.
XX
AC      ABM00939;
XX
DT      15-JAN-2004 (first entry)
XX
DE      Diacosoma sp. red fluorescent protein (RFP), K70M.
XX
KM      Fluorescent protein; resonance energy transfer; pH; detection;
KM      red fluorescent protein; RFP; mutant; muteln.
XX
OS      Diacosoma sp.
OS      Synthetic.
XX
FH      Key          Location/Qualifiers
FT      Misc-difference 70
FT      /note= "Wild-type Lys is substituted with Met"
XX
FN      US2003170911-A1.
XX
PD      11-SEP-2003.
XX
PF      26-FEB-2001; 2001US-00794308.
PR      26-FEB-2001; 2001US-00794308.
XX
PA      (TSIE/) TSIE R Y.
PA      (ZACH/) ZACHARIAS D A.
PA      (BAIR/) BAIRD G S.
XX
PI      Tsien RY, Zacharias DA, Baird GS;
XX
DR      WPI; 2003-802418/75.
XX
PT      Fluorescent proteins containing a mutation that reduces or eliminates its
PT      ability to oligomerize which gives more reliable fluorescence resonance
PT      energy transfer results and are useful to detect molecule interaction,
PT      enzymes, or sample pH.
XX
PS      Example 2; Page; 0pp; English.
XX
CC      The invention relates to a non-oligomerizing fluorescent protein
CC      containing a mutation that reduces or eliminates its ability to
CC      oligomerize. The fluorescent protein gives more reliable fluorescence
CC      resonance energy transfer results and are useful to detect molecule
CC      interaction, enzymes, or sample pH. These are also used to identify
CC      agents or conditions that regulate expression of control sequences. The
CC      present sequence is Diacosoma sp. red fluorescent protein (RFP) mutant.
CC      Note: This sequence is not shown in the specification, however this
CC      sequence is constructed based on Diacosoma sp. wild-type RFP protein
CC      shown in page 30-31 (ABM00918)
XX
SQ      Sequence 225 AA;

Query Match          99.5%; Score 1208; DB 7; Length 225;
Best Local Similarity 99.6%; Pred. No. 4,9e-127;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  MRSSKNVKEPFRFVMEGTNGHGFETEGSGRPRYEGHNTVLTXTKGGRPLFPADI 60
DB      1  MRSSKNVKEPFRFVMEGTNGHGFETEGSGRPRYEGHNTVLTXTKGGRPLFPADI 60
QY      61  LSPQGYGSKVYVKHPADI PDYKKLSPEGFKMERVMNFEDGGVTVTQDSLQDGCFTY 120

```

```

DB      61  LSPQGYGSKVYVKHPADI PDYKKLSPEGFKMERVMNFEDGGVTVTQDSLQDGCFTY 120
QY      121  KYKFIGVNPSPDGPVMQKTKMGWASTERTLYPRDGLKGEIHKALKLKDGGHYLVFPSI 180
DB      121  KYKFIGVNPSPDGPVMQKTKMGWASTERTLYPRDGLKGEIHKALKLKDGGHYLVFPSI 180

RESULT 42
ADH34500
ID      ADH34500 standard; protein; 225 AA.
XX
AC      ADH34500;
XX
DT      11-MAR-2004 (first entry)
XX
DE      Diacosoma sp. Dared mutant S148A/S203A.
XX
KM      Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;
KM      Chidnarian; Anthozoan; labeling; colouring agents; pigment;
KM      analyte detection assay; selectable marker; sunscreen; selective filter;
KM      fluorescence resonance energy transfer; FRRT; biosensor;
KM      whole cell marker; second messenger detector; in vivo marker;
KM      fluorescence activated cell sorting; fluorescent timer;
KM      red fluorescent protein; Dared; mutant; muteln.
XX
OS      Synthetic.
OS      Diacosoma sp.
XX
FH      Key          Location/Qualifiers
FT      Misc-difference 146
FT      /note= "Ala replaces wild-type Ser. Corresponds to GFP
FT      residue 148"
FT      Misc-difference 197
FT      /note= "Ala replaces wild-type Ser. Corresponds to GFP
FT      residue 203"
XX
FN      WO2003057833-A2.
XX
PD      17-JUL-2003.
XX
PF      23-DEC-2002; 2002WO-US041418.
PR      26-DEC-2001; 2001US-0343128P.
XX
PA      (CLON-) CLONTECH LAB INC.
XX
PI      Bulina ME, Chudakov D, Lukyanov KA;
XX
DR      WPI; 2003-607998/57.
XX
PT      Novel nucleic acid encoding interconverted mutant of chromo-or
PT      fluorescent protein which are useful as biosensors, coloring agents.
XX
PS      Example 1; Page; 56pp; English.
XX
CC      The invention relates to interconverted mutants of chromoproteins (CP) or
CC      fluorescent proteins (FP) and nucleic acids encoding them. The mutant is
CC      derived from a Chidnarian species, preferably a non-bioluminescent
CC      Chidnarian species, and most preferably an Anthozoan species. The
CC      invention is based on the finding that although green fluorescent protein
CC      (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of
CC      homology, there are certain positions (referred to as 148, 165, 167 and
CC      203; numbering corresponds to GFP) that are occupied by noticeably
CC      different residues in the two types of proteins. Mutagenesis of the
CC      residues in these key positions in, for example, a fluorescent protein,
CC      to those found in a chromoprotein is therefore proposed to confer
CC      chromoprotein activity on the fluorescent protein mutant, with
CC      chromoproteins being able to be converted into fluorescent proteins in a

```

CC similar manner. The invention also relates to expression constructs, vectors, host cells and host cell progeny comprising a nucleic acid of the invention; the recombinant production of an interconverted chromoprotein or fluorescent protein mutant; and antibodies specific for interconverted mutant proteins of the invention. The interconverted mutants are useful in any application that employs a chromoprotein or fluorescent protein. Fluorescent protein mutants having chromoprotein activity can be useful as colouring agents in, for example, food compositions, pharmaceuticals, cosmetics and living organisms. Proteins with chromoprotein activity are also useful as labels in biological assays, as selectable markers in recombinant DNA applications (e.g. the production of transgenic cells and organisms), and are also useful as screens and selective filters. Chromoproteins are also useful as fluorescent protein activity filters in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, as markers of whole cells to detect changes in multicellular reorganisation and migration, as second messenger detectors, as in vivo markers in animals (e.g., transgenic animals), in fluorescence activated cell sorting applications, in protease cleavage assays, and in assays to determine the phospholipid composition in biological membranes. Proteins with fluorescent protein activity can also be used as fluorescent timers, where the switch of one fluorescent colour to another (e.g., green to red) is concomitant with the ageing of the protein and is useful for determination of the activation or deactivation of gene expression. The present sequence represents a Discosoma sp. red fluorescent protein DsRed mutant generated in an example of the invention. The present sequence is not shown in the specification, but was derived from the wild-type DsRed sequence (ADH34469) shown in Fig 1 and the information provided on page 42.

CC Sequence 225 AA;

Query Match 99.5%; Score 1208; DB 7; Length 225;  
Best Local Similarity 99.1%; Pred. No. 4, 9e-127;  
Matches 223; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSKRVIKRPMFKVMRMEGTGNGHEFEIIEGEGRPYEGHNTVKLKVTGKGPLPRAWDI 60  
DB 1 MRSKRVIKRPMFKVMRMEGTGNGHEFEIIEGEGRPYEGHNTVKLKVTGKGPLPRAWDI 60  
QY 61 LSPQFQYGSKVYVYKHADIPDYKLSFPEGFKMERVMNFDGCVAVTVDSSIQDGCFTY 120  
DB 61 LSPQFQYGSKVYVYKHADIPDYKLSFPEGFKMERVMNFDGCVAVTVDSSIQDGCFTY 120  
QY 121 KYVFIGVNFPSDGPVWQKTMGWEASTERLYPRDGVKGIHNAKXKDGSHLVVERKSI 180  
DB 121 KYVFIGVNFPSDGPVWQKTMGWEASTERLYPRDGVKGIHNAKXKDGSHLVVERKSI 180  
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQERTGRRHHLFL 225  
DB 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQERTGRRHHLFL 225

RESULT 43

ID AAE28921 standard; protein; 225 AA.

AC AAE28921;

DB 27-DEC-2002 (first entry)

XX Discosoma sp. drFP583 (NFP-6) mutant protein, E5 (V105A+S197T).

KM Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KM fluorescence activated cell sorting application; fluorescent timer;  
KM biosensor; fluorescence resonance energy transfer application; FRET;  
KM colouring agent; recombinant DNA application; analyte detection assay;  
KM sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
KM mutain.

XX Discosoma sp.  
OS Synthetic.  
XX

FN Key Location/Qualifiers

FT Misc-difference 105 /note= "Wild-type Val substituted with Ala"

FT Misc-difference 217 /note= "Wild-type Ser substituted with Thr"

PN WO200268459-A2.

XX 06-SEP-2002.

PD 20-FEB-2002; 2002WO-US005749.

PF 21-FEB-2001; 2001US-0270983P.

PR 04-DEC-2001; 2001US-00006922.

PA (CLON-) CLONTECH LAB INC.

PI Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Fradkov A;

XX WPI; 2002-691654/74.

PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant of an aggregating Cnidarian chromo- or fluorescent protein or mutant for analyte detection assays or fluorescence activated cell sorting applications.

XX Disclosure; Page; 80pp; English.

XX The invention relates to nucleic acid molecules encoding non-aggregating chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are useful in analyte detection assays, as colouring agents, as markers in recombinant DNA applications, as screens or filters, in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, in screening assays, as second messenger detectors, in fluorescence activated cell sorting applications, in protease cleavage assays or as fluorescent timers. The present sequence is Discosoma sp. drFP583 (NFP-6) mutant protein of the invention. Note: This sequence is not shown in the specification, but is derived from Discosoma sp. drFP583 (NFP-6) wild-type protein shown as SEQ ID NO:8 (AAE28833) in page 70-71 of the specification

XX Sequence 225 AA;

Query Match 99.4%; Score 1207; DB 5; Length 225;  
Best Local Similarity 99.1%; Pred. No. 6, 3e-127;  
Matches 223; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSKRVIKRPMFKVMRMEGTGNGHEFEIIEGEGRPYEGHNTVKLKVTGKGPLPRAWDI 60  
DB 1 MRSKRVIKRPMFKVMRMEGTGNGHEFEIIEGEGRPYEGHNTVKLKVTGKGPLPRAWDI 60  
QY 61 LSPQFQYGSKVYVYKHADIPDYKLSFPEGFKMERVMNFDGCVAVTVDSSIQDGCFTY 120  
DB 61 LSPQFQYGSKVYVYKHADIPDYKLSFPEGFKMERVMNFDGCVAVTVDSSIQDGCFTY 120  
QY 121 KYVFIGVNFPSDGPVWQKTMGWEASTERLYPRDGVKGIHNAKXKDGSHLVVERKSI 180  
DB 121 KYVFIGVNFPSDGPVWQKTMGWEASTERLYPRDGVKGIHNAKXKDGSHLVVERKSI 180  
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQERTGRRHHLFL 225  
DB 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQERTGRRHHLFL 225

RESULT 44

ID AAE17541 standard; protein; 225 AA.

AC AAE17541;

DB 22-APR-2002 (first entry)

XX Discosoma sp. fluorescent protein E5.

KX	Fluorescent timer protein; protein movement; translocation; trafficking;
KM	promoter activity; gene expression; transgenic plant; gene modification;
KW	protein age; E5.
OS	Diccosoma sp.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 105 /note= "Wild type Val substituted with Ala"
FT	Misc-difference 197 /note= "Wild type Ser substituted with Thr"
PT	
PN	MO200196373-A2.
XX	
PD	20-DEC-2001.
XX	
PF	13-JUN-2001; 2001WO-US019097.
PR	14-JUN-2000; 2000US-0211607P.
XX	
PA	(CLON-) CLONTECH LAB INC.
XX	
PI	Fradekov AF, Terakikh A;
XX	
DR	WI; 2002-154595/20.
XX	N-PSDB; AAD28208.
PT	New fluorescent timer proteins comprising an emission spectrum that changes over time from a first wavelength to a second wavelength, useful for monitoring intracellular protein movement, translocation, trafficking or stability.
PT	
XX	
S8	Claim 5; Fig 2; 89pp; English.

Query Match	99.4%	Score 1207;	DB 5;	Length 225;
Best Local Similarity	99.1%;	Pred. No. 6.3e-127;		
Matches 223;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	MRSSKNVTKEMRFKVRMEGTVNGHGFETEGEGEGRPYEGHNTVTLKVTGKGGLPFADI	60	
Db	1	MRSSKNVTKEMRFKVRMEGTVNGHGFETEGEGEGRPYEGHNTVTLKVTGKGGLPFADI	60	
QY	61	LSPOFGYSKTVYKHPADIPIYKKLSPEEGFKMERVNNPFGDGVTVTQDSSLQDCTIY	120	

Db	61	LSPOQVSKYVYVHPALIPDYKLSFPBEGFWMEKVMNPFEDGVATVWQDSLQDSCFLY	120
QY	121	KYKFLGNVPSBDGPMQKKTGMGEMSTERLYPRDGLGGEIKALKLRGGHYLYEFSK	180
Db	121	KYKFLGNVPSBDGPMQKKTGMGEMSTERLYPRDGLGGEIKALKLRGGHYLYEFSK	180
QY	181	YMAKPPVQLPGYYVDSKLDITSNHDYTVYEQYERTBSGHHLLF	225
Db	181	YMAKPPVQLPGYYVDSKLDITSNHDYTVYEQYERTBSGHHLLF	225

RESULT 45	
ADY51734	
ID	ADY51734 standard; protein; 225 AA.
XX	
AC	ADY51734;
XX	
DT	05-MAY-2005 (first entry)
XX	
DE	Discosoma sp. red fluorescent protein (RFP) mutant I125R.
XX	
KW	fluorescence; mutagenesis; red fluorescent protein; protein interaction; mte1n.
XX	
OS	Discosoma sp.
OS	Synthetic.
XX	
PH	Key
FT	Misc-difference 125
FT	/note= "Wild type Ile substituted by Arg"

PT New non-oligomerizing fluorescent protein containing at least one  
 PT mutation that reduces or eliminates the ability of the protein to  
 PT oligomerize, useful for making better and new assays for molecular  
 PT biology.  
 XX  
 PS  
 PS Claim 11; Page; 117pp; English.  
 XX  
 CC This invention relates to a novel non-oligomerizing fluorescent protein.  
 CC Specifically, it refers to the presence of at least one mutation in the  
 CC fluorescent protein that reduces or eliminates the ability of the protein  
 CC to oligomerize. The present invention describes fluorescent proteins  
 CC derived from naturally occurring green or red fluorescent proteins and  
 CC provides a fusion protein that comprises a non-oligomerizing fluorescent  
 CC protein linked to at least one protein of interest. As such, these fusion  
 CC proteins can be used in methods and compositions to determine the pH of a  
 CC sample, or whether the sample contains an enzyme, molecule or agent that  
 CC regulates the activity of an expression control sequence. Furthermore,  
 CC they may be used to identify a specific interaction of molecules, such  
 CC that they are useful for improving or developing new assays in the field  
 CC of molecular biology. This polypeptide sequence is the Discosoma sp. red  
 CC fluorescent protein (RFP) mutant of the invention. NOTE: This sequence is  
 CC not given in the invention but it is derived from seqid 12 and  
 CC information provided in the claims.

Query Match 99.4%; Score 1207; DB 5; Length 225;



Best Local Similarity 99.6%; Pred. No. 6.3e-127;  
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSKNVIEFNRFRKVRNEGTVNGHGFELIEGEGRPGYGHNTVTKLTKGGLPFPAMD 60  
1 MSSKNVIEFNRFRKVRNEGTVNGHGFELIEGEGRPGYGHNTVTKLTKGGLPFPAMD 60  
Db 1 MSSKNVIEFNRFRKVRNEGTVNGHGFELIEGEGRPGYGHNTVTKLTKGGLPFPAMD 60

QY 61 LSPFOYGSKVVYKVPADIPDYKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120  
61 LSPFOYGSKVVYKVPADIPDYKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120  
Db 61 LSPFOYGSKVVYKVPADIPDYKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120

QY 121 KVKFIVGNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
121 KVKFIVGNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
Db 121 KVKFIVGNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 46  
ABW00941  
ID ABW00941 standard; protein; 225 AA.  
XX  
AC ABW00941;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein (RFP), I125R.  
XX  
KW Fluorescent protein; resonance energy transfer; pH; detection;  
KM red fluorescent protein; RFP; mutant; mutcin.  
XX  
OS Discosoma sp.  
OS Synthetic.  
XX  
Key Location/Qualifiers  
FH Misc-difference 125  
FT /note= "Wild-type Ile is substituted with Arg"  
XX  
PN US2003170911-A1.  
XX  
PD 11-SEP-2003.  
XX  
PF 26-FEB-2001; 2001US-00794308.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
XX  
PA (TSIE/) TSIE R Y.  
PA (ZACH/) ZACHARIAS D A.  
PA (BAIR/) BAIRD G S.  
XX  
PI Tsien RY, Zacharias DA, Baird GS;  
XX  
DR WPI; 2003-802418/75.  
XX  
PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX  
Example 3; Page; Opp; English.  
XX  
The invention relates to a non-oligomerizing fluorescent protein  
XX containing a mutation that reduces or eliminates its ability to  
XX oligomerize. The fluorescent protein gives more reliable fluorescence  
XX resonance energy transfer results and are useful to detect molecule  
XX interaction, enzymes, or sample pH. These are also used to identify  
XX agents or conditions that regulate expression of control sequences. The  
XX present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
XX Note: This sequence is not shown in the specification, however this  
XX sequence is constructed based on Discosoma sp. wild-type RFP protein  
XX shown in page 30-31 (ABW00918)

XX SQ Sequence 225 AA;  
SQ Query Match 99.4%; Score 1207; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 6.3e-127;  
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSKNVIEFNRFRKVRNEGTVNGHGFELIEGEGRPGYGHNTVTKLTKGGLPFPAMD 60  
1 MSSKNVIEFNRFRKVRNEGTVNGHGFELIEGEGRPGYGHNTVTKLTKGGLPFPAMD 60  
Db 1 MSSKNVIEFNRFRKVRNEGTVNGHGFELIEGEGRPGYGHNTVTKLTKGGLPFPAMD 60

QY 61 LSPFOYGSKVVYKVPADIPDYKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120  
61 LSPFOYGSKVVYKVPADIPDYKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120  
Db 61 LSPFOYGSKVVYKVPADIPDYKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120

QY 121 KVKFIVGNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
121 KVKFIVGNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
Db 121 KVKFIVGNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 47  
ABW00940  
ID ABW00940 standard; protein; 225 AA.  
XX  
AC ABW00940;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein (RFP), I125K.  
XX  
KW Fluorescent protein; resonance energy transfer; pH; detection;  
KM red fluorescent protein; RFP; mutant; mutcin.  
XX  
OS Discosoma sp.  
OS Synthetic.  
XX  
Key Location/Qualifiers  
FH Misc-difference 125  
FT /note= "Wild-type Ile is substituted with Lys"  
XX  
PN US2003170911-A1.  
XX  
PD 11-SEP-2003.  
XX  
PF 26-FEB-2001; 2001US-00794308.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
XX  
PA (TSIE/) TSIE R Y.  
PA (ZACH/) ZACHARIAS D A.  
PA (BAIR/) BAIRD G S.  
XX  
PI Tsien RY, Zacharias DA, Baird GS;  
XX  
DR WPI; 2003-802418/75.  
XX  
PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX  
Example 3; Page; Opp; English.  
XX  
The invention relates to a non-oligomerizing fluorescent protein  
XX containing a mutation that reduces or eliminates its ability to  
XX oligomerize. The fluorescent protein gives more reliable fluorescence  
XX resonance energy transfer results and are useful to detect molecule  
XX interaction, enzymes, or sample pH. These are also used to identify  
XX agents or conditions that regulate expression of control sequences. The

CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
CC shown in page 30-31 (ABW00918)

XX Sequence 225 AA;

Query Match 99.4%; Score 1207; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 6.3e-127;  
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVKEKEMRFRKMEGTVNGHFEIIEGEGRPYEGHNTVYKLYTKGGRPLPFANDI 60  
DB 1 MRSSKNVKEKEMRFRKMEGTVNGHFEIIEGEGRPYEGHNTVYKLYTKGGRPLPFANDI 60  
QY 61 LSPQFQYGSKYVYVGHADIPDYKKLSFPEGFKMERVMNMFEDGVVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVYVGHADIPDYKKLSFPEGFKMERVMNMFEDGVVTVTQDSSLQDGCFTY 120  
QY 121 KYKFRGVNFPESDGPVMQKKTGMWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
DB 121 KYKFRGVNFPESDGPVMQKKTGMWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 48  
ADL46222  
ID ADL46222 standard; protein; 225 AA.  
XX  
AC ADL46222;

DT 20-MAY-2004 (first entry)  
XX

DE Discosoma red fluorescent (DsRed) protein with 1125r mutation.

XX red fluorescent protein; DsRed; fluorescence; red wavelength.  
KM oligomerization; tetramerization; immunoassay; hybridization assay.

XX Discosoma sp.

OS WO2003086446-A1.

PN 23-OCT-2003.

PD 09-APR-2003; 2003WO-US010879.

PF 10-APR-2002; 2002US-00121258.

PR 29-JUL-2002; 2002US-00209208.

PA (REGC ) UNIV CALIFORNIA.

PI Tsien RY, Campbell RE, Baird GS;

XX WPI; 2003-845265/78.

PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.

XX Disclousure; SEQ ID NO 20; 166p; English.

XX The invention relates to a polynucleotide sequence encoding a Discosoma  
CC red fluorescent protein (DsRed) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced

CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC protein with an 1125r mutation.

XX Sequence 225 AA;

Query Match 99.4%; Score 1207; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 6.3e-127;  
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVKEKEMRFRKMEGTVNGHFEIIEGEGRPYEGHNTVYKLYTKGGRPLPFANDI 60  
DB 1 MRSSKNVKEKEMRFRKMEGTVNGHFEIIEGEGRPYEGHNTVYKLYTKGGRPLPFANDI 60  
QY 61 LSPQFQYGSKYVYVGHADIPDYKKLSFPEGFKMERVMNMFEDGVVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVYVGHADIPDYKKLSFPEGFKMERVMNMFEDGVVTVTQDSSLQDGCFTY 120  
QY 121 KYKFRGVNFPESDGPVMQKKTGMWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
DB 121 KYKFRGVNFPESDGPVMQKKTGMWEASTERLYPRDGVLKGEIHKALKDKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 49  
ABW00933  
ID ABW00933 standard; protein; 225 AA.  
XX  
AC ABW00933;

DT 15-JAN-2004 (first entry)  
XX

DE Discosoma sp. red fluorescent protein (RFP), K83F.

XX Fluorescent protein; resonance energy transfer; pH; detection;

KM red fluorescent protein; RFP; mutant; mutain.

XX Discosoma sp.

OS Synthetic.

OS Key Location/Qualifiers

FT Misc-difference 83 /note= "Wild-type Lys is substituted with Phe"

XX US2003170911-A1.

XX 11-SEP-2003.

XX 26-FEB-2001; 2001US-00794308.

XX 26-FEB-2001; 2001US-00794308.

XX (TSIE/) TSIE R Y.

XX (ZACH/) ZACHARIAS D A.

XX (BAIR/) BAIRD G S.

XX Tsien RY, Zacharias DA, Baird GS;

XX WPI; 2003-802418/75.

XX Fluorescent proteins containing a mutation that reduces or eliminates its

XX ability to oligomerize which gives more reliable fluorescence resonance

XX energy transfer results and are useful to detect molecule interaction,

XX enzymes, or sample pH.

XX Example 2; Page; Opp; English.

XX The invention relates to a non-oligomerising fluorescent protein

CC containing a mutation that reduces or eliminates its ability to  
 CC oligomerize. The fluorescent protein gives more reliable fluorescence  
 CC resonance energy transfer results and are useful to detect molecule  
 CC interaction, enzymes, or sample pH. These are also used to identify  
 CC agents or conditions that regulate expression of control sequences. The  
 CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
 CC Note: This sequence is not shown in the specification, however this  
 CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
 CC shown in page 30-31 (ABW00918)

XX Sequence 225 AA;

Query Match 99.3%; Score 1206; DB 7; Length 225;

Best Local Similarity 99.6%; Pred. No. 8.2e-127;  
 Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVRMEGTNGHGFIEGEGRPYGHNTVYKTKTGGPLPPAMD 60  
 DB 1 MSSKNVKEFMRFKVRMEGTNGHGFIEGEGRPYGHNTVYKTKTGGPLPPAMD 60  
 QY 61 LSPQFYSGKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGAVTVTQDSSLQDGCFTY 120  
 DB 61 LSPQFYSGKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGAVTVTQDSSLQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVWQKKTMGMEASTERTLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
 DB 121 KYKFIGVNPSPDGPVWQKKTMGMEASTERTLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
 QY 181 YNAKKPVQLPGYVVYDSKLDITSNEDYTVVEQYERTGRHHLFL 225  
 DB 181 YNAKKPVQLPGYVVYDSKLDITSNEDYTVVEQYERTGRHHLFL 225

RESULT 50

ABW00934 standard; protein; 225 AA.

AC ABW00934;

DT 15-JAN-2004 (first entry)

XX Discosoma sp. red fluorescent protein (RFP), K83W.

XX Fluorescent protein; resonance energy transfer; pH; detection;

KM red fluorescent protein; RFP; mutant; mutcin.

XX Discosoma sp.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 83 /note= "Wild-type Lys is substituted with Trp"

XX US2003170911-A1.

XX 11-SEP-2003.

XX 26-FEB-2001; 2001US-00794308.

XX 26-FEB-2001; 2001US-00794308.

XX (TSIE/) TSIE R Y.

XX (ZACH/) ZACHARIAS D A.

XX (BAIR/) BAIRD G S.

XX Tsien RY, Zacharias DA, Baird GS;

XX WPI, 2003-802418/75.

XX Fluorescent proteins containing a mutation that reduces or eliminates its  
 PT ability to oligomerize which gives more reliable fluorescence resonance  
 PT energy transfer results and are useful to detect molecule interaction,  
 PT enzymes, or sample pH.

XX Example 2; Page; 0pp; English.

PS The invention relates to a non-oligomerizing fluorescent protein  
 XX containing a mutation that reduces or eliminates its ability to  
 CC oligomerize. The fluorescent protein gives more reliable fluorescence  
 CC resonance energy transfer results and are useful to detect molecule  
 CC interaction, enzymes, or sample pH. These are also used to identify  
 CC agents or conditions that regulate expression of control sequences. The  
 CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
 CC Note: This sequence is not shown in the specification, however this  
 CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
 CC shown in page 30-31 (ABW00918)

XX Sequence 225 AA;

Query Match 99.3%; Score 1206; DB 7; Length 225;

Best Local Similarity 99.6%; Pred. No. 8.2e-127;  
 Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVRMEGTNGHGFIEGEGRPYGHNTVYKTKTGGPLPPAMD 60  
 DB 1 MSSKNVKEFMRFKVRMEGTNGHGFIEGEGRPYGHNTVYKTKTGGPLPPAMD 60  
 QY 61 LSPQFYSGKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGAVTVTQDSSLQDGCFTY 120  
 DB 61 LSPQFYSGKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGAVTVTQDSSLQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVWQKKTMGMEASTERTLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
 DB 121 KYKFIGVNPSPDGPVWQKKTMGMEASTERTLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
 QY 181 YNAKKPVQLPGYVVYDSKLDITSNEDYTVVEQYERTGRHHLFL 225  
 DB 181 YNAKKPVQLPGYVVYDSKLDITSNEDYTVVEQYERTGRHHLFL 225

RESULT 51

ADH34501 standard; protein; 225 AA.

AC ADH34501;

DT 11-MAR-2004 (first entry)

XX Discosoma sp. Dared mutant S148A/K167M.

XX Chromoprotein; fluorescent protein; CP; PP; interconverted mutant;

KM Chidarian; Anthozoa; labelling; colouring agents; pigment;

KM analytical detection assay; selectable marker; biosensor; selective filter;

KM fluorescence resonance energy transfer; FRET; biosensor;

KM whole cell marker; second messenger detector; in vivo marker;

KM fluorescence activated cell sorting; fluorescent timer;

XX red fluorescent protein; Dared; mutant; mutcin.

OS Synthetic.

XX Discosoma sp.

XX Key Location/Qualifiers

FT Misc-difference 146 /note= "Ala replaces wild-type Ser. Corresponds to GFP

FT residue 148"

FT Misc-difference 163 /note= "Met replaces wild-type Lys. Corresponds to GFP

XX residue 167"

XX WO2003057833-A2.

XX 17-JUL-2003.

XX 23-DEC-2002; 2002WO-US041418.

XX 26-DEC-2001; 2001US-0343128P.

XX (CLON-) CLONTECH LAB INC.  
 PA  
 XX  
 XX Bulina ME, Chudakov D, Lukyanov KA;  
 FI  
 DR WPI, 2003-607998/57.  
 XX  
 XX  
 PT Novel nucleic acid encoding interconverted mutant of chromo-or  
 fluorescent protein which are useful as biosensors, coloring agents.  
 PS  
 XX Example 1; Page; 56pp; English.

CC The invention relates to interconverted mutants of chromoproteins (CP) or  
 CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
 CC derived from a Chitridian species, preferably a non-bioluminescent  
 CC Chitridian species, and most preferably an Anchozoan species. The  
 CC invention is based on the finding that although green fluorescent protein  
 CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
 CC homology, there are certain positions (referred to as 148, 165, 167 and  
 CC 203; numbering corresponds to GFP) that are occupied by noticeably  
 CC different residues in the two types of proteins. Mutagenesis of the  
 CC residues in these key positions in, for example, a fluorescent protein,  
 CC to those found in a chromoprotein is therefore proposed to confer  
 CC chromoprotein activity on the fluorescent protein mutant, with  
 CC chromoproteins being able to be converted into fluorescent proteins in a  
 CC similar manner. The invention also relates to expression constructs,  
 CC vectors, host cells and host cell progeny comprising a nucleic acid of  
 CC the invention; the recombinant production of an interconverted  
 CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
 CC interconverted mutant proteins of the invention. The interconverted  
 CC mutants are useful in any application that employs a chromoprotein or  
 CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
 CC activity can useful as colouring agents in, for example, food  
 CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
 CC with chromoprotein activity are also useful as labels in biological  
 CC analyte detection assays, as selectable markers in recombinant DNA  
 CC applications (e.g. the production of transgenic cells and organisms), and  
 CC are also useful as sunscreens and selective filters. Chromoprotein  
 CC mutants having fluorescent protein activity useful in fluorescence  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
 CC changes in multicellular reorganisation and migration, as second  
 CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
 CC animals), in fluorescence activated cell sorting applications, in  
 CC protease cleavage assays, and in assays to determine the phospholipid  
 CC composition in biological membranes. Proteins with fluorescent protein  
 CC activity can also be used as fluorescent timers, where the switch of one  
 CC fluorescent colour to another (e.g., green to red) is concomitant with  
 CC the ageing of the protein and is useful for determination of the  
 CC activation or deactivation of gene expression. The present sequence  
 CC represents a *Discosoma* sp. red fluorescent protein Dared mutant generated  
 CC in an example of the invention. The present sequence is not shown in the  
 CC specification, but was derived from the wild-type Dared sequence  
 CC (ADH34489) shown in Fig 1 and the information provided on page 42.

XX  
 XX Sequence 225 AA;

Query Match 99.3%; Score 1205; DB 7; Length 225;  
 Best Local Similarity 99.1%; Pred. No. 1.1e-126;  
 Matches 223; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSSKNVKEPFRFVRMEGTNGHBEFEISGEGRPRVGHNTVTLKATKGGPLPFANDI 60  
 DB 1 MRSSKNVKEPFRFVRMEGTNGHBEFEISGEGRPRVGHNTVTLKATKGGPLPFANDI 60  
 OY 61 LSPQOYGSKVYKVPADIPDYKLSFPEGFKMERVMFEDGVTVTYQDSSLQDGCFTY 120  
 DB 61 LSPQOYGSKVYKVPADIPDYKLSFPEGFKMERVMFEDGVTVTYQDSSLQDGCFTY 120  
 OY 121 KYKFGVNPSPDGPVWOKKTGMESTERLYPRDGLKGEIHAKLKDGGHYLVEFKSI 180  
 DB 121 KYKFGVNPSPDGPVWOKKTGMESTERLYPRDGLKGEIHAKLKDGGHYLVEFKSI 180  
 OY 121 KYKFGVNPSPDGPVWOKKTGMESTERLYPRDGLKGEIHAKLKDGGHYLVEFKSI 180  
 DB 121 KYKFGVNPSPDGPVWOKKTGMESTERLYPRDGLKGEIHAKLKDGGHYLVEFKSI 180

OY 181 YMAKPPVQLPGSYVYVDSKLDITSHNEDTYIVEQYTERGHRHLFL 225  
 DB 181 YMAKPPVQLPGSYVYVDSKLDITSHNEDTYIVEQYTERGHRHLFL 225

RESULT 52  
 ID ABP56678 standard; protein; 240 AA.  
 AC ABP56678;  
 DT 25-MAR-2003 (first entry)

DE Mammalian codon optimised *Discosoma* red fluorescent protein.

KW Mammalian codon optimised *Discosoma* red fluorescent protein; *Discosoma*;  
 KM red fluorescent protein; directed evolution.

OS *Discosoma* sp.  
 OS Mammalia.  
 OS Synthetic.

FT Key Location/Qualifiers

FT Misc-difference 24 /note= "possible mutations at this point are N24S, N24R  
 and N24H"

FT Misc-difference 125 /note= "possible mutations at this point are F125L and  
 F125V"

FT Misc-difference 164 /note= "a possible mutation at this point is K164M"

FT Misc-difference 183 /note= "a possible mutation at this point is M183K"

FT Misc-difference 206 /note= "encoded by AAC"

FT WO200294992-A2.

PD 28-NOV-2002.

PF 20-MAY-2002; 2002WO-US015968.

PR 18-MAY-2001; 2001US-0291871P.

PA (RIGB-) RIGEL PHARM INC.

XX Peelie B;

DR WPI; 2003-120798/11.

DR N-PSDB; AB222476.

PT New *Discosoma* red fluorescent protein, useful for functional screens as a  
 reporter for gene transcription, for target characterization and  
 localization of fusion proteins, or for scaffolds for protein and peptide  
 libraries.

PS Claim 1; Fig 1; 22pp; English.

CC The present invention describes an isolated *Discosoma* red fluorescent  
 CC protein (1) comprising a 241 residue amino acid sequence (see ABP56678,  
 CC S1), with one or more point mutations at amino acid position N24, F125,  
 CC K164, or M183. Also described: (1) a fusion protein comprising (1); (2)  
 CC an isolated nucleic acid encoding (1); (3) a vector comprising the  
 CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
 CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
 CC encoding (1); and (6) methods of making a fluorescent variant. The  
 CC *Discosoma* red fluorescent proteins are useful for functional screens as a  
 CC reporter for gene transcription (e.g. as a fusion protein), for target  
 CC characterisation and localisation of fusion proteins, or for scaffolds  
 CC for protein and peptide libraries. The fluorescent proteins can also be  
 CC used as selectable markers or reporter molecules for a variety of  
 CC assays, including methods that use fluorescence activated cell sorting  
 CC (FACS) as a selection mechanism. The method of directed protein evolution

is useful for obtaining improved variants of red fluorescent protein. The variants of Discosoma red fluorescent protein have greatly improved brightness, expression, and/or folding kinetics as compared to wild type or a codon optimized variant. The present sequence represents a mammalian codon optimized Discosoma red fluorescent protein, which is used in an example from the present invention

SQ Sequence 240 AA;

Query Match 99.3%; Score 1205; DB 6; Length 240;  
Best Local Similarity 99.1%; Pred. No. 1.2e-126;  
Matches 223; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 MSSKNVKEFMRFKVMEGTNGHFEIEGEGRPYGHNTVTKLKTGGPLPRAWMI 60  
2 VASSKNVKEFMRFKVMEGTNGHFEIEGEGRPYGHNTVTKLKTGGPLPRAWMI 61  
61 LSPQFGSKAVVKKPADIPDYKLSFPGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
62 LSPQFGSKAVVKKPADIPDYKLSFPGFKMERVMNFDGCVTVTODSSLQDGCFTY 121  
121 KKKFIGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLXDGHYLVEFKSI 180  
122 KKKFIGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLXDGHYLVEFKSI 181  
181 YNAKKRVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225  
182 YNAKKRVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 226

RESULT 53

AAE28923  
ID AAE28923 standard; protein; 225 AA.

27-DEC-2002 (first entry)  
Discosoma sp. drfp583 (NFP-6) mutant protein, B83 (N42H+V71A+I180H).

Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
fluorescence activated cell sorting application; fluorescent timer;  
biosensor; fluorescence resonance energy transfer application; FRET;  
colouring agent; recombinant DNA application; analyte detection assay;  
sunscreen; second messenger detector; drfp583 protein; NFP-6; mutant;  
mutain.

XX Discosoma sp.  
OS Synthetic.

XX Key Location/Qualifiers

FT MISC-difference 42 /note= "Wild-type Asn substituted with His"  
FT MISC-difference 71 /note= "Wild-type Val substituted with Ala"  
FT MISC-difference 180 /note= "Wild-type Ile substituted with Val"  
XX MO200268459-A2.

XX 06-SEP-2002.

XX 20-FEB-2002; 2002MO-US005749.

XX 21-FEB-2001; 2001US-0270983P.  
XX 04-DEC-2001; 2001US-00006922.

XX (CLON-) CLONTECH LAB INC.

XX Lukyanov S, Lukyanov K, Yarnushevich Y, Savitsky A, Fradkov A;  
XX WPI, 2002-691654/74.  
XX

PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Chidatrian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.

XX Disclosure; Page: 80pp; English.

XX The invention relates to nucleic acid molecules encoding non-aggregating  
XX chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
XX useful in analyte detection assays, as colouring agents, as markers in  
XX recombinant DNA applications, as sunscreens or filters, in fluorescence  
XX resonance energy transfer (FRET) applications, as biosensors in  
XX prokaryotic and eukaryotic cells, in screening assays, as second  
XX messenger detectors, in fluorescence activated cell sorting applications,  
XX in protease cleavage assays or as fluorescent timers. The present  
XX sequence is Discosoma sp. drfp583 (NFP-6) mutant protein of the  
XX invention. Note: This sequence is not shown in the specification, but is  
XX derived from Discosoma sp. drfp583 (NFP-6) wild-type protein shown as SEQ  
XX ID NO:8 (AAE28833) in page 70-71 of the specification

SQ Sequence 225 AA;

Query Match 99.2%; Score 1204; DB 5; Length 225;  
Best Local Similarity 98.7%; Pred. No. 1.4e-126;  
Matches 222; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MSSKNVKEFMRFKVMEGTNGHFEIEGEGRPYGHNTVTKLKTGGPLPRAWMI 60  
1 MSSKNVKEFMRFKVMEGTNGHFEIEGEGRPYGHNTVTKLKTGGPLPRAWMI 60  
61 LSPQFGSKAVVKKPADIPDYKLSFPGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
61 LSPQFGSKAVVKKPADIPDYKLSFPGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
121 KKKFIGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLXDGHYLVEFKSI 180  
121 KKKFIGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLXDGHYLVEFKSI 180  
181 YNAKKRVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225  
181 YNAKKRVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225

RESULT 54

AAE28925  
ID AAE28925 standard; protein; 225 AA.

27-DEC-2002 (first entry)  
Discosoma sp. drfp583 (NFP-6) mutant protein, A64 (V71M+V105A+S197T).

Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
fluorescence activated cell sorting application; fluorescent timer;  
biosensor; fluorescence resonance energy transfer application; FRET;  
colouring agent; recombinant DNA application; analyte detection assay;  
sunscreen; second messenger detector; drfp583 protein; NFP-6; mutant;  
mutain.

XX Discosoma sp.

OS Synthetic.

XX Key Location/Qualifiers

FT MISC-difference 71 /note= "Wild-type Val substituted with Met"  
FT MISC-difference 105 /note= "Wild-type Val substituted with Ala"  
FT MISC-difference 197 /note= "Wild-type Ser substituted with Thr"  
XX MO200268459-A2.  
XX

```
PD 06-SEP-2002.
XX 20-FEB-2002; 2002MO-US005749.
XX 21-FEB-2001; 2001US-0270983P.
XX 04-DEC-2001; 2001US-00006922.
XX (CLON-) CLONTECH LAB INC.
XX Lukanov S, Lukanov K, Yarushevich Y, Savitsky A, Fradkov A;
XX WPI; 2002-691654/74.
XX
XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant
XX of an aggregating Chidarian chromo- or fluorescent protein or mutant for
XX analyte detection assays or fluorescence activated cell sorting
XX applications.
XX
XX Discloure; Page; 80pp; English.
XX
XX The invention relates to nucleic acid molecules encoding non-aggregating
XX chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are
XX useful in analyte detection assays, as colouring agents, as markers in
XX recombinant DNA applications, as sunscreens or filters, in fluorescence
XX resonance energy transfer (FRET) applications, as biosensors in
XX prokaryotic and eukaryotic cells, in screening assays, as second
XX messenger detectors, in fluorescence activated cell sorting applications,
XX in protease cleavage assays or as fluorescent timers. The present
XX sequence is Discosoma sp. drfp583 (NFP-6) mutant protein of the
XX invention. Note: This sequence is not shown in the specification, but is
XX derived from Discosoma sp. drfp583 (NFP-6) wild-type protein shown as SEQ
XX ID NO:8 (AAE28833) in page 70-71 of the specification
XX
XX Sequence 225 AA;
XX
XX Query Match 99.2%; Score 1204; DB 5; Length 225;
XX Best Local Similarity 98.7%; Pred. No. 1.4e-126;
XX Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MRSSKNVYKEFRMRKRVMEGTVNGHGFELIEBGGGRPRYEGNHYTKLKTGKGPPLPFANDI 60
XX 1 MRSSKNVYKEFRMRKRVMEGTVNGHGFELIEBGGGRPRYEGNHYTKLKTGKGPPLPFANDI 60
XX 1 LSPQOYGSKYVYVHPADIPYKLLSPFEGFKMRVWVNFEDGVAITYVQDSSLQDGCFTY 120
XX 61 LSPQOYGSKYVYVHPADIPYKLLSPFEGFKMRVWVNFEDGVAITYVQDSSLQDGCFTY 120
XX 61 LSPQOYGSKYVYVHPADIPYKLLSPFEGFKMRVWVNFEDGVAITYVQDSSLQDGCFTY 120
XX
XX 121 KVFIFGVNFPSPDGPVMQKTMGWEASTERLYPRDGVLGKEIHKALKLDGGHYLVEFKSI 180
XX 121 KVFIFGVNFPSPDGPVMQKTMGWEASTERLYPRDGVLGKEIHKALKLDGGHYLVEFKSI 180
XX
XX 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTYVEQYERTBGRHHLFL 225
XX 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTYVEQYERTBGRHHLFL 225
XX
XX RESULT 55
XX AAE28924
XX AAE28924 standard; protein; 225 AA.
XX
XX AAE28924;
XX
XX 27-DEC-2002 (first entry)
XX
XX Discosoma sp. drfp583 (NFP-6) mutant protein, B57 (V105A+I161T+S197A).
XX
XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;
XX fluorescence activated cell sorting application; fluorescent timer;
XX biosensor; fluorescence resonance energy transfer application; FRET;
XX colouring agent; recombinant DNA application; analyte detection assay;
XX sunscreen; second messenger detector; drfp583 protein; NFP-6; mutant;
XX mutein.
XX
XX
```

```
OS Discosoma sp.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX FT Misc-difference 105 /note= "Wild-type Val substituted with Ala"
XX FT Misc-difference 161 /note= "Wild-type Ile substituted with Thr"
XX FT Misc-difference 197 /note= "Wild-type Ser substituted with Ala"
XX
XX WO200268459-A2.
XX
XX 06-SEP-2002.
XX
XX 20-FEB-2002; 2002MO-US005749.
XX
XX 21-FEB-2001; 2001US-0270983P.
XX 04-DEC-2001; 2001US-00006922.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Lukanov S, Lukanov K, Yarushevich Y, Savitsky A, Fradkov A;
XX WPI; 2002-691654/74.
XX
XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant
XX of an aggregating Chidarian chromo- or fluorescent protein or mutant for
XX analyte detection assays or fluorescence activated cell sorting
XX applications.
XX
XX Discloure; Page; 80pp; English.
XX
XX The invention relates to nucleic acid molecules encoding non-aggregating
XX chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are
XX useful in analyte detection assays, as colouring agents, as markers in
XX recombinant DNA applications, as sunscreens or filters, in fluorescence
XX resonance energy transfer (FRET) applications, as biosensors in
XX prokaryotic and eukaryotic cells, in screening assays, as second
XX messenger detectors, in fluorescence activated cell sorting applications,
XX in protease cleavage assays or as fluorescent timers. The present
XX sequence is Discosoma sp. drfp583 (NFP-6) mutant protein of the
XX invention. Note: This sequence is not shown in the specification, but is
XX derived from Discosoma sp. drfp583 (NFP-6) wild-type protein shown as SEQ
XX ID NO:8 (AAE28833) in page 70-71 of the specification
XX
XX Sequence 225 AA;
XX
XX Query Match 99.0%; Score 1202; DB 5; Length 225;
XX Best Local Similarity 98.7%; Pred. No. 2.3e-126;
XX Matches 222; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 MRSSKNVYKEFRMRKRVMEGTVNGHGFELIEBGGGRPRYEGNHYTKLKTGKGPPLPFANDI 60
XX 1 MRSSKNVYKEFRMRKRVMEGTVNGHGFELIEBGGGRPRYEGNHYTKLKTGKGPPLPFANDI 60
XX
XX 61 LSPQOYGSKYVYVHPADIPYKLLSPFEGFKMRVWVNFEDGVAITYVQDSSLQDGCFTY 120
XX 61 LSPQOYGSKYVYVHPADIPYKLLSPFEGFKMRVWVNFEDGVAITYVQDSSLQDGCFTY 120
XX 61 LSPQOYGSKYVYVHPADIPYKLLSPFEGFKMRVWVNFEDGVAITYVQDSSLQDGCFTY 120
XX
XX 121 KVFIFGVNFPSPDGPVMQKTMGWEASTERLYPRDGVLGKEIHKALKLDGGHYLVEFKSI 180
XX 121 KVFIFGVNFPSPDGPVMQKTMGWEASTERLYPRDGVLGKEIHKALKLDGGHYLVEFKSI 180
XX
XX 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTYVEQYERTBGRHHLFL 225
XX 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTYVEQYERTBGRHHLFL 225
XX
XX RESULT 56
XX ADH34503
XX ADH34503 standard; protein; 225 AA.
XX
```

AC ADH34503;  
XX  
DT 11-MAR-2004 (first entry)  
DE Discosoma sp. Darked mutant S148A/K167M/S203A.  
XX  
XX Chromoprotein; fluorescent protein; CP; PP; interconverted mutant;  
XX Cnidarian; Anthozoan; labelling; colouring agents; pigment;  
XX analyte detection assay; selectable marker; sunscreen; selective filter;  
XX fluorescence resonance energy transfer; FRET; biosensor;  
XX whole cell marker; second messenger detector; in vivo marker;  
XX red fluorescent protein; Darked; mutant; mutain.  
XX  
OS Synthetic.  
XX Discosoma sp.  
XX  
XX Key Location/Qualifiers  
XX Misc-difference 146 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
XX FT residue 148"  
XX FT Misc-difference 161 /note= "Ser replaces wild-type Ile. Corresponds to GFP  
XX FT residue 165"  
XX FT Misc-difference 197 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
XX FT residue 203"  
XX WO2003057833-A2.  
XX  
XX 17-JUL-2003.  
XX PD  
XX 23-DEC-2002; 2002WO-US041418.  
XX PF  
XX 26-DEC-2001; 2001US-0343128P.  
XX PR  
XX (CLON-) CLONTECH LAB INC.  
XX PA  
XX Bulina ME, Chudakov D, Lukyanov KA;  
XX PI  
XX WPI; 2003-607998/57.  
XX DR  
XX Novel nucleic acid encoding interconverted mutant of chromo-or  
XX PT fluorescent protein which are useful as biosensors, coloring agents.  
XX  
XX Example 1; Page: 56pp; English.  
XX  
XX The invention relates to interconverted mutants of chromoproteins (CP) or  
XX CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
XX CC derived from a Cnidarian species, preferably a non-bioluminescent  
XX CC Cnidarian species, and most preferably an Anthozoan species. The  
XX CC invention is based on the finding that although green fluorescent protein  
XX CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
XX CC homology, there are certain positions (referred to as 148, 165, 167 and  
XX CC 203; numbering corresponds to GFP) that are occupied by noticeably  
XX CC different residues in the two types of proteins. Mutagenesis of the  
XX CC residues in these key positions in, for example, a fluorescent protein,  
XX CC to those found in a chromoprotein is therefore proposed to confer  
XX CC chromoprotein activity on the fluorescent protein mutant, with  
XX CC chromoproteins being able to be converted into fluorescent proteins in a  
XX CC similar manner. The invention also relates to expression constructs,  
XX CC vectors, host cells and host cell progeny comprising a nucleic acid of  
XX CC the invention; the recombinant production of an interconverted  
XX CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
XX CC interconverted mutant proteins of the invention. The interconverted  
XX CC mutants are useful in any application that employs a chromoprotein or  
XX CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
XX CC activity can be useful as colouring agents in, for example, food  
XX CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
XX CC with chromoprotein activity are also useful as labels in biological  
XX CC analyte detection assays, as selectable markers in recombinant DNA  
XX CC applications (e.g. the production of transgenic cells and organisms), and  
XX CC are also useful as sunscreens and selective filters. Chromoprotein

CC mutants having fluorescent protein activity useful in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
CC changes in multicellular reorganisation and migration, as second  
CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
CC animals), in fluorescence activated cell sorting applications, in  
CC protease cleavage assays, and in assays to determine the phospholipid  
CC composition in biological membranes. Proteins with fluorescent protein  
CC activity can also be used as fluorescent timers, where the switch of one  
CC fluorescent colour to another (e.g., green to red) is concomitant with  
CC the ageing of the protein and is useful for determination of the  
CC activation or deactivation of gene expression. The present sequence  
CC represents a Discosoma sp. red fluorescent protein Darked mutant generated  
CC in an example of the invention. The present sequence is not shown in the  
CC specification, but was derived from the wild-type Darked sequence  
CC (ADH34489) shown in Fig 1 and the information provided on page 42.  
XX  
XX Sequence 225 AA;  
SQ  
XX  
XX Query Match 99.0%; Score 1202; DB 7; Length 225;  
XX Best Local Similarity 98.7%; Pred. No. 2,3e-12e;  
XX Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MRSSKNVTKEMRPFYRMGGTVNGHFEPIEGEGRPYEGHNTVTKLYTKGGPLPFAMDI 60  
DB 1 MRSSKNVTKEMRPFYRMGGTVNGHFEPIEGEGRPYEGHNTVTKLYTKGGPLPFAMDI 60  
QY 61 LSPQFQYSKVVYKHPADIPDYKLSFPBGFKMERVMNPEEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYSKVVYKHPADIPDYKLSFPBGFKMERVMNPEEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFTGVNPPSDGPMOKKTMGEMASTERLYPRDGLVKEITKAKLKDGGHYLVPEFSI 180  
DB 121 KYKFTGVNPPSDGPMOKKTMGEMASTERLYPRDGLVKEISHKALKLKDGGHYLVPEFSI 180  
QY 181 YMAKKPVOLPGYVYVDSKLDITSNEDYTYVEQYRTREGHHFL 225  
DB 181 YMAKKPVOLPGYVYVDAKLDITSNEDYTYVEQYRTREGHHFL 225  
RESULT 57  
ADH34502  
ID ADH34502 standard; protein; 225 AA.  
XX  
XX ADH34502;  
XX  
XX DT 11-MAR-2004 (first entry)  
XX  
XX Discosoma sp. Darked mutant S148A/K167M/S203A.  
XX  
XX Chromoprotein; fluorescent protein; CP; PP; interconverted mutant;  
XX CC Cnidarian; Anthozoan; labelling; colouring agents; pigment;  
XX CC analyte detection assay; selectable marker; sunscreen; selective filter;  
XX CC fluorescence resonance energy transfer; FRET; biosensor;  
XX CC whole cell marker; second messenger detector; in vivo marker;  
XX CC fluorescence activated cell sorting; fluorescent timer;  
XX CC red fluorescent protein; Darked; mutant; mutain.  
XX  
OS Synthetic.  
XX Discosoma sp.  
XX  
XX Key Location/Qualifiers  
XX FT Misc-difference 146 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
XX FT residue 148"  
XX FT Misc-difference 163 /note= "Ser replaces wild-type Lys. Corresponds to GFP  
XX FT residue 167"  
XX FT Misc-difference 197 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
XX FT residue 203"  
XX WO2003057833-A2.

XX 17-JUL-2003.  
XD 23-DEC-2002; 2002MO-US041418.  
PF 26-DEC-2001; 2001US-0343128P.  
XX (CLON-) CLONTECH LAB INC.  
PA Bulina ME, Chudakov D, Lukyanov KA;  
XX WPI; 2003-607998/57.  
XX Novel nucleic acid encoding interconverted mutant of chromo-or  
PT fluorescent protein which are useful as biosensors, coloring agents.  
XX  
XX Example 1; Page; 56pp; English.  
PS The invention relates to interconverted mutants of chromoproteins (CP) or  
CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
CC derived from a Chudakov species, preferably a non-bioluminescent  
CC Chudakov species, and most preferably an Anthozoan species. The  
CC invention is based on the finding that although green fluorescent protein  
CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
CC homology, there are certain positions (referred to as 148, 165, 167 and  
CC 203; numbering corresponds to GFP) that are occupied by noticeably  
CC different residues in the two types of proteins. Mutagenesis of the  
CC residues in these key positions in, for example, a fluorescent protein,  
CC to those found in a chromoprotein is therefore proposed to confer  
CC chromoprotein activity on the fluorescent protein mutant, with  
CC a similar manner. The invention also relates to expression constructs in a  
CC vectors, host cells and host cell progeny comprising a nucleic acid of  
CC the invention, the recombinant production of an interconverted  
CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
CC interconverted mutant proteins of the invention. The interconverted  
CC mutants are useful in any application that employs a chromoprotein or  
CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
CC activity can be useful as coloring agents in, for example, food  
CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
CC with chromoprotein activity are also useful as labels in biological  
CC analytical detection assays, as selectable markers in recombinant DNA  
CC applications (e.g., the production of transgenic cells and organisms), and  
CC are also useful as sensors and selective filters. Chromoprotein  
CC mutants having fluorescent protein activity useful in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
CC changes in multicellular reorganization and migration, as second  
CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
CC animals), in fluorescence activated cell sorting applications, in  
CC protease cleavage assays, and in assays to determine the phospholipid  
CC composition in biological membranes. Proteins with fluorescent protein  
CC activity can also be used as fluorescent timers, where the switch of one  
CC fluorescent colour to another (e.g., green to red) is concomitant with  
CC the aging of the protein and is useful for determination of the  
CC activation or deactivation of gene expression. The present sequence  
CC represents a Discosoma sp. red fluorescent protein DsRed mutant generated  
CC in an example of the invention. The present sequence is not shown in the  
CC specification, but was derived from the wild-type DsRed sequence  
CC (ADH34489) shown in Fig 1 and the information provided on page 42.  
XX  
XX Sequence 225 AA;  
SQ

Query Match 99.0%; Score 1202; DB 7; Length 225;  
Best Local Similarity 98.7%; Pred. No. 2.3e-126;  
Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSKNVKEPFRFVMEGTNGHBFETEGSGRPRVGHNTVLKATKGGPLFPADI 60  
DB 1 MSSKNVKEPFRFVMEGTNGHBFETEGSGRPRVGHNTVLKATKGGPLFPADI 60  
QY 61 LSPQFQYGSKVYKHPADIIPDYKLLSPFGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
DB 121 KYKPIGVNFPESDGYVMQKTKMGWEASTERLLYPRGVLKGEITHKLKLDGGHYLVERFSI 180  
DB 121 KYKPIGVNFPESDGYVMQKTKMGWEASTERLLYPRGVLKGEITHKLKLDGGHYLVERFSI 180  
QY 181 YMAKKPVOLPGYVVYVDSKLDITSHNEDYTYIYOYERTEGRHHLFL 225  
DB 181 YMAKKPVOLPGYVVYVDAKLDITSHNEDYTYIYOYERTEGRHHLFL 225

DB 61 LSPQFQYGSKVYKHPADIIPDYKLLSPFGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
QY 121 KYKPIGVNFPESDGYVMQKTKMGWEASTERLLYPRGVLKGEITHKLKLDGGHYLVERFSI 180  
DB 121 KYKPIGVNFPESDGYVMQKTKMGWEASTERLLYPRGVLKGEITHKLKLDGGHYLVERFSI 180  
QY 181 YMAKKPVOLPGYVVYVDSKLDITSHNEDYTYIYOYERTEGRHHLFL 225  
DB 181 YMAKKPVOLPGYVVYVDAKLDITSHNEDYTYIYOYERTEGRHHLFL 225

RESULT 58  
ABP56685  
ID ABP56685 standard; protein; 240 AA.  
XX  
XX ABP56685;  
XX  
XX 25-MAR-2003 (first entry)  
XX  
XX DE Discosoma red fluorescent protein variant K93R.  
XX  
XX Mammalian codon optimised Discosoma red fluorescent protein; variant;  
XX Discosoma; red fluorescent protein; directed evolution; mutant.  
XX  
XX Discosoma sp.  
XX Synthetic.  
XX WO200294992-A2.  
XX  
XX 28-NOV-2002.  
XX  
XX 20-MAY-2002; 2002MO-US015968.  
XX  
XX 18-MAY-2001; 2001US-0291871P.  
XX (RIGE-) RIGEL PHARM INC.  
XX  
XX Peelie B;  
XX WPI; 2003-120798/11.  
XX  
XX New Discosoma red fluorescent protein, useful for functional screens as a  
XX reporter for gene transcription, for target characterization and  
XX localization of fusion proteins, or for scaffolds for protein and peptide  
XX libraries.  
XX  
XX Claim 17; Page; 22pp; English.  
XX  
XX The present invention describes an isolated Discosoma red fluorescent  
XX protein (1) comprising a 241 residue amino acid sequence (see ABP56678,  
XX S1), with one or more point mutations at amino acid position N24, R125,  
XX K164, or M183. Also described: (1) a fusion protein comprising (1); (2)  
XX an isolated nucleic acid encoding (1); (3) a vector comprising the  
XX nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
XX retroviral cDNA expression library comprising the nucleic acid of (2), or  
XX encoding (1); and (6) methods of making a fluorescent variant. The  
XX Discosoma red fluorescent proteins are useful for functional screens as a  
XX reporter for gene transcription (e.g. as a fusion protein), for target  
XX characterization and localization of fusion proteins, or for scaffolds  
XX for protein and peptide libraries. The fluorescent proteins can also be  
XX used as selectable markers or reporter molecules for a variety of  
XX assays, including methods that use fluorescence activated cell sorting  
XX (FACS) as a selection mechanism. The method of directed protein evolution  
XX is useful for obtaining improved variants of red fluorescent protein. The  
XX variants of Discosoma red fluorescent protein have greatly improved  
XX brightness, expression, and/or folding kinetics as compared to wild type  
XX or a codon optimised variant. The present sequence represents a mammalian  
XX codon optimised Discosoma red fluorescent protein variant from the  
XX present invention. N.B. The present sequence is not given in the  
XX specification, but is derived from the protein in Fig 1 as stated in the  
XX claims  
SQ Sequence 240 AA;





21-FEB-2001; 2001US-0270983p.  
04-DEC-2001; 2001US-00006922.  
(CLON-) CLONTECH LAB INC.  
Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Fradkov A;  
WPI; 2002-691654/74.  
New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
analyte detection assays or fluorescence activated cell sorting  
applications.  
Disclosure; Page; 80pp; English.  
The invention relates to nucleic acid molecules encoding non-aggregating  
chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
useful in analyte detection assays, as colouring agents, as markers in  
recombinant DNA applications, as screens or filters, in fluorescence  
resonance energy transfer (FRET) applications, as biosensors in  
prokaryotic and eukaryotic cells, in screening assays, as second  
messenger detectors, in fluorescence activated cell sorting applications,  
in protease cleavage assays or as fluorescent timers. The present  
sequence is *Discooma* sp. drFP583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the type protein shown as SEQ  
ID No:8 (AAE28833) in page 70-71 of the specification  
Sequence 225 AA;

Query Match	98.8%;	Score 1199;	DB 5;	Length 225;
Best Local Similarity	98.2%;	Pred. No. 5e-126;		
Matches 221;	Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0
QY	1	MRSSKNVIEKFNRFKRMGMGTGNGHGFELGEGEGRPVEGHNTVYKLKVTKGGLPFPAMD	60	
Db	1	MRSSKNVIEKFNRFKRMGMGTGNGHGFELGEGEGRPVEGHNTVYKLKVTKGGLPFPAMD	60	
QY	61	LSDFQYGSKVVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVVYVTDSSLDGCFIY	120	
Db	61	LSDFQYGSKVVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVVYVTDSSLDGCFIY	120	
QY	121	KXKPEIGVNPFSOGPNVQKKTGMGEASTELLYRPDGVKGEIHNKLYKLDGSHYLYVERKSI	180	
Db	121	KXKPEIGVNPFSOGPNVQKKTGMGEASTELLYRPDGVKGEIHNKLYKLDGSHYLYVERKSI	180	
QY	181	YMAKKPQVLPQGVYVYVDSKLDITSNHEDYIVVEQYERTEGRHHLFL	225	
Db	181	YMAKKPQVLPQGVYVYVDTKLDITSNHEDYIVVEQYERTEGRHHLFL	225	

  

RESULT 61
ID ADH34505
ADH34505 standard; protein; 225 AA.
XX
AC ADH34505;
XX
DT 11-MAR-2004 (first entry)
XX
DE Discosoma sp. D6red mutant S148C/I165N/S203A.
XX
KW Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;
KW Chlmdharian; Anthozoan; labelling; colouring agents; pigment;
KW analytical detection assay; selectable marker; sunscreen; selective filter;
KW fluorescence resonance energy transfer; FRET; biosensor;
KW whole cell marker; second messenger detector; in vivo marker;
KW fluorescence activated cell sorting; fluorescent timer;
KW red fluorescent protein; D6red; mutant; mutcin.
XX
XX Synthetic.
OS Discosoma sp.

Key	Location/Qualifiers
Misc-difference 146	/note="Cys replaces wild-type Ser. Corresponds to GFP residue 148"
Misc-difference 161	/note="Asn replaces wild-type Ile. Corresponds to GFP residue 165"
Misc-difference 197	/note="Ala replaces wild-type Ser. Corresponds to GFP residue 203"
WO2003057833-A2.	
17-JUL-2003.	
23-DEC-2002; 2002MO-US041418.	
26-DEC-2001; 2001US-0343128P.	
(CLON-) CLONTECH LAB INC.	
Bulina ME, Chudakov D, Lukyanov KA;	
WPI; 2003-607998/57.	
Novel nucleic acid encoding interconverted mutant of chromo-or fluorescent protein which are useful as biosensors, coloring agents.	
Example 1; Page; 56pp; English.	
The invention relates to interconverted mutants of chromoproteins (CP) or fluorescent proteins (FP) and nucleic acids encoding them. The mutant is derived from a Chitmanian species, preferably a non-bioluminescent Chitmanian species, and most preferably an Anthozoan species. The invention is based on the finding that although green fluorescent protein (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of homology, there are certain positions (referred to as 148, 165, 167 and 203; numbering corresponds to GFP) that are occupied by noticeably different residues in the two types of proteins. Mutagenesis of the residues in these key positions in, for example, a fluorescent protein, to those found in a chromoprotein is therefore proposed to confer chromoprotein activity on the fluorescent protein mutant, with chromoproteins being able to be converted into fluorescent proteins in a similar manner. The invention also relates to expression constructs, vectors, host cells and host cell progeny comprising a nucleic acid of the invention; the recombinant production of an interconverted chromoprotein or fluorescent protein mutant; and antibodies specific for interconverted mutant proteins of the invention. The interconverted mutants are useful in any application that employs a chromoprotein or fluorescent protein. Fluorescent protein mutants having chromoprotein activity can be useful as colouring agents in, for example, food compositions, pharmaceuticals, cosmetics and living organisms. Proteins with chromoprotein activity are also useful as labels in biological analyte detection assays, as selectable markers in recombinant DNA applications (e.g. the production of transgenic cells and organisms), and are also useful as screens and selective filters. Chromoprotein mutants having fluorescent protein activity useful in biosensors in resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, as markers of whole cells to detect changes in multicellular reorganisation and migration, as second messenger detectors, as in vivo markers in animals (e.g., transgenic animals), in fluorescence activated cell sorting applications, in protease cleavage assays, and in assays to determine the phospholipid composition in biological membranes. Proteins with fluorescent protein activity can also be used as fluorescent timers, where the switch of one fluorescent colour to another (e.g., green to red) is concomitant with the ageing of the protein and is useful for determination of the activation or deactivation of gene expression. The present sequence represents a <i>Discosoma</i> sp. red fluorescent protein Dared mutant generated in an example of the invention. The present sequence is not shown in the specification but was derived from the wild-type Dared sequence (ADH31489) shown in Fig 1 and the information provided on page 42.	

XX Sequence 225 AA;  
SQ  
Query Match 98.8%; Score 1199; DB 7; Length 225;  
Best Local Similarity 98.7%; Pred. No. 5e-126; Mismatches 1; Indels 0; Gaps 0;  
Matches 222; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MRSKNVKEFMRFKVMEGTNGHGFELIEGEGRPYEGHNTVTLKVTKGGLPFANDI 60  
DB 1 MRSKNVKEFMRFKVMEGTNGHGFELIEGEGRPYEGHNTVTLKVTKGGLPFANDI 60  
QY 61 LSPQFQYGSKYVVKAPADIPDYKLSFPEGFKMERVMNFDGGVTVTTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVVKAPADIPDYKLSFPEGFKMERVMNFDGGVTVTTQDSSLQDGCFTY 120  
QY 121 KVKFTGVNPPSDGPMQKKTWGMESTERLYPRDGLKGEIHKALKKDGGHYLVFESKI 180  
DB 121 KVKFTGVNPPSDGPMQKKTWGMESTERLYPRDGLKGEIHKALKKDGGHYLVFESKI 180  
QY 181 YNAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225  
DB 181 YNAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225  
RESULT 62  
ABP56684  
ID ABP56684 standard; protein; 240 AA.  
XX ABP56684;  
XX 25-MAR-2003 (first entry)  
XX Discosoma red fluorescent protein variant F125L.  
XX Mammalian codon optimised Discosoma red fluorescent protein; variant;  
XX Discosoma; red fluorescent protein; directed evolution; mutant.  
XX Discosoma sp.  
XX Synthetic.  
XX WO200294992-A2.  
XX 28-NOV-2002.  
XX 20-MAY-2002; 2002WO-US015968.  
XX 18-MAY-2001; 2001US-0291871P.  
XX (RIGR-) RIGEL PHARM INC.  
XX Peelle B;  
XX WPI; 2003-120798/11.  
XX New Discosoma red fluorescent protein, useful for functional screens as a  
XX reporter for gene transcription, for target characterization and  
XX localization of fusion proteins, or for scaffolds for protein and peptide  
XX libraries.  
XX Claim 15; Page; 22pp; English.  
XX The present invention describes an isolated Discosoma red fluorescent  
XX protein (I) comprising a 241 residue amino acid sequence (see ABP56678,  
XX S1), with one or more point mutations at amino acid position N24, F125,  
XX K164, or M163. Also described: (1) a fusion protein comprising (1); (2)  
XX an isolated nucleic acid encoding (1); (3) a vector comprising the  
XX nucleic acid of (2); (4) a host cell comprising the nucleic acid of (2), or  
XX retroviral cDNA expression library comprising the nucleic acid of (2), or  
XX encoding (1); and (6) methods of making a fluorescent variant. The  
XX Discosoma red fluorescent proteins are useful for functional screens as a  
XX reporter for gene transcription (e.g. as a fusion protein), for target  
XX characterisation and localisation of fusion proteins, or for scaffolds  
XX for protein and peptide libraries. The fluorescent proteins can also be

CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of Discosoma red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised Discosoma red fluorescent protein variant from the  
CC present invention. N.B. The present sequence is not given in the  
CC specification, but is derived from the protein in Fig 1 as stated in the  
CC claims  
XX  
SQ Sequence 240 AA;  
Query Match 98.8%; Score 1199; DB 6; Length 240;  
Best Local Similarity 98.7%; Pred. No. 5e-126; Mismatches 1; Indels 0; Gaps 0;  
Matches 222; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MRSKNVKEFMRFKVMEGTNGHGFELIEGEGRPYEGHNTVTLKVTKGGLPFANDI 60  
DB 2 VRSKNVKEFMRFKVMEGTNGHGFELIEGEGRPYEGHNTVTLKVTKGGLPFANDI 61  
QY 61 LSPQFQYGSKYVVKAPADIPDYKLSFPEGFKMERVMNFDGGVTVTTQDSSLQDGCFTY 120  
DB 62 LSPQFQYGSKYVVKAPADIPDYKLSFPEGFKMERVMNFDGGVTVTTQDSSLQDGCFTY 121  
QY 121 KVKFTGVNPPSDGPMQKKTWGMESTERLYPRDGLKGEIHKALKKDGGHYLVFESKI 180  
DB 122 KVKFTGVNPPSDGPMQKKTWGMESTERLYPRDGLKGEIHKALKKDGGHYLVFESKI 181  
QY 181 YNAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 225  
DB 182 YNAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTBGRHHLFL 226  
RESULT 63  
ABP56683  
ID ABP56683 standard; protein; 240 AA.  
XX ABP56683;  
XX 25-MAR-2003 (first entry)  
XX Discosoma red fluorescent protein variant K164M.  
XX Discosoma red fluorescent protein; variant;  
XX Mammalian codon optimised Discosoma red fluorescent protein; variant;  
XX Discosoma; red fluorescent protein; directed evolution; mutant.  
XX Discosoma sp.  
XX Synthetic.  
XX WO200294992-A2.  
XX 28-NOV-2002.  
XX 20-MAY-2002; 2002WO-US015968.  
XX 18-MAY-2001; 2001US-0291871P.  
XX (RIGR-) RIGEL PHARM INC.  
XX Peelle B;  
XX WPI; 2003-120798/11.  
XX New Discosoma red fluorescent protein, useful for functional screens as a  
XX reporter for gene transcription, for target characterization and  
XX localization of fusion proteins, or for scaffolds for protein and peptide  
XX libraries.  
XX Claim 14; Page; 22pp; English.  
XX The present invention describes an isolated Discosoma red fluorescent



	Best Local Similarity	98.2%;	Pred. No. 1.1e-125;	Mismatches	221;	Conservative	2;	Mismatches	2;	Indels	0;	Gaps	0.
Qy	1	MRSSKNVYKEPMRKRVMEGVNNGHEFEIIEEGEGRPYEGHNTVLKVTXGGPLPFPADI	60										
Dd	1	MRSSKNVIKEMRKRVMEGVNNGHEFEIIEEGESRPVEGHNTVLKVTXGGPLPFPADI	60										
Qy	61	LSPQFGSKVYVGHFADIPDYKKLSPEEGFKMERVMNFEDGCVTTVTQDSLDGCTTY	120										
Dd	61	LSPQFGSKVYVGHFADIPDYKKLSPEEGFKMERVMNFEDGCVTTVTQDSSLDGCPHY	120										
Qy	121	KVKRTGVNFPEDSGPVMOCKTKMGWEASTERLKPVSGLVKGELHKALKKDGGHYLFPSKI	180										
Dd	121	KVKRTGVNFPEDSGPVMOCKTKMGWEASTERLRPRDSVLGSESMALKLXDGGHYLFPSKI	180										
Qy	181	YMAKKPVQLPGYYVYDSTKLDTTSHNEDYTIYEQYERTGRRHLFL	225										
Dd	181	YMAKKPVQLPGYYIVDAKDLTTSNEDYTIYEQYERTGRRHLFL	225										

RESULT 65  
AD078068  
ID AD078068 standard; protein; 236 AA.  
XX  
AC ADO78068;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
Dd Corallimorpharia red fluorescent protein mutant Red1.  
XX  
Dd Corallimorpharia red fluorescent protein; green fluorescent protein;  
XX gene expression detection; regeneration; multiple labeling system;  
XX in vivo marker; microinjection assay; taxonomic marker; colour indicator;  
XX food additive; cosmetic; Actinodiscus; Discosoma; muteln; mutant; Red1;  
XX coral-like anemone.  
XX  
OS Corallimorpharia.  
XX  
PN US2004110225-A1.  
XX  
PD 10-JUN-2004.  
XX  
PF 09-DEC-2002; 2002US-00314936.  
XX  
PR 09-DEC-2002; 2002US-00314936.  
XX  
PA (GIBB/) GIBBS P D L.  
PA (CART/) CARTER R W.  
PA (SCHM/) SCHMALE M C.  
PI Glbbs PDL, Carter RW, Schmale MC;  
XX  
WPI 2004-467659/44.  
DR N-PSDB; ADO78067.  
XX  
PT New first mutant fluorescent protein having an optical property  
PT relatively different to a corresponding optical property produced by a  
PT wild-type red or green fluorescent protein, useful for detecting  
PT expression of a gene.  
XX  
PS Claim 6; SEQ ID NO 2; 30pp; English.  
XX  
CC The invention describes a first mutant fluorescent protein (I) comprising  
CC an amino acid sequence that differs from that of a corresponding wild-  
CC type red fluorescent protein by an amino acid substitution, where the  
CC first mutation fluorescent protein mutant has an optical property that  
CC differs relative to the corresponding optical property produced by the  
CC wild-type red or green fluorescent protein. (I) is useful for detecting  
CC expression of gene. The polynucleotide (II) encoding (I) is useful for  
CC detecting expression of a gene which involves introducing (II) into a  
CC cell or organism, allowing the cell to replicate, and detecting  
CC expression of the nucleic acid by emission of fluorescent light. The  
CC expression of the nucleic acid expression is detected in vivo or in

	CC	vltro-(ii) is useful for regenerating an animal such as a zebrafish
	CC	which contains (iii). (I) is useful as a marker for detecting expression
	CC	of a gene, in biochemical assays, and as reagents. (II) is also useful in
	CC	multiple labeling systems, as in vivo markers such as in mRNA
	CC	microinjection assays, and as taxonomic markers for studies of genetics,
	CC	colour indicators in diagnostic kits, coloured food additives, and
	CC	cosmetic ingredients. (II) is useful in research for up or down
	CC	regulation, to monitor promoter activity, to allow longer term monitoring
	CC	and to localise proteins. (I) has enhanced properties such as
	CC	substantially enhanced fluorescence and reduced toxicity. This is the
	CC	amino acid sequence of Redi, a mutant of the red fluorescent protein
	CC	isolated from an aquatic species believed to be either an Actinodiscus or
	CC	Discosoma species.
SQ		Sequence 236 AA:
Query Match	98.5%; Score 1196; DB 8; Length 236;	
Best Local Similarity	98.2%; Pred. No. 1.2e-125;	
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;		
Oy	1 MRSSKNVTKEMRKVKVMETGVNCHERRIRBEGGRPYEENNTYKLVTKGSPLPFAMD I 60	
Db	1 MSCSKNYIKERMRPKVMETGVNCHERRIBEBGGRPPEGNNTVKLVTKGSLPFAMD I 60	
Oy	61 LSPFOYSKKYYVGHADIPDKLSPPEGPKMERVMNFEDGVVVYTODSSLDDGCFT Y 120	
Db	61 LSPFOYSKKYYVGHADIPDKLSPPGPKMERVMNFEDGSVVYTODSLLDGCFT Y 120	
Oy	121 KYKITGVNFPDSGVMOCKTMGWASTERTLYPRDGLVKEIHKLKLDKGSHYLVEFKSI 180	
Db	121 KYKITGVNFPDSGVMOCKTMGWASTERTLYPRDGLVKEIHKLKLDKGSHYLVEFKTI 180	
Oy	181 YMAPPVOLPGYYYVDSDITTSHNEDITYEOYRTTREGRRHFL 225	
Db	181 YMAPPVOLPGYYYVDSDIDITSNKDYTYVOYRTEGRRHFL 225	
RESULT 66		
ADO78070	ID ADO78070 standard; protein; 236 AA.	
XX AC	ADO78070;	
XX DT	26-AUG-2004 (first entry)	
XX DE	CoralImmorpharia red fluorescent protein mutant Redti.	
XX KW	fluorescent protein; red fluorescent protein; green fluorescent protein;	
KW	gene expression detection; regeneration; multiple labeling system;	
KW	in vivo marker; microinjection assay; taxonomic marker; colour indicator;	
KW	food additive; cosmetic; Actinodiscus; Discosoma; mutein; mutant; Redti;	
KW	coral-like anemone.	
OS	CoralImmorpharia.	
PN	US2004110225-A1.	
XX PD	10-JUN-2004.	
XX PF	09-DEC-2002; 2002US-00314936.	
XX PR	09-DEC-2002; 2002US-00314936.	
PA PA	(GIBB/) GIBBS P D L.	
PA	(CART/) CARTER R W.	
PI PI	(SCHM/) SCHMALE M C.	
Gibbs PDL, Carter RW, Schmale MC;		
DR N-PADB; ADO78069.		
New first mutant fluorescent protein having an optical property		

PT relatively different to a corresponding optical property produced by a  
 PT wild-type red or green fluorescent protein, useful for detecting  
 expression of a gene.

Claim 7, SEQ ID NO 4; 30pp; English.

The invention describes a first mutant fluorescent protein (I) comprising an amino acid sequence that differs from that of a corresponding wild-type red fluorescent protein by an amino acid substitution, where the first mutation fluorescent protein mutant has an optical property that differs relative to the corresponding optical property produced by the wild-type red or green fluorescent protein. (I) is useful for detecting expression of gene. The polynucleotide (II) encoding (I) is useful for detecting expression of a gene which involves introducing (II) into a cell or organism, allowing the cell to replicate, and detecting expression of the nucleic acid by emission of fluorescent light. The expression of the nucleic acid expression is detected in vivo or in vitro. (II) is useful for regenerating an animal such as a zebrafish which contains (II). (I) is useful as a marker for detecting expression of a gene, in biochemical assays, and as reagents. (I) is also useful in multiple labeling systems, as in vivo markers such as in mRNA microinjection assays, and as taxonomic markers for studies of genetics, colour indicators in diagnostic kits, coloured food additives, and cosmetics ingredients. (I) is useful in research for up or down regulation, to monitor promoter activity, to allow longer term monitoring and to localise proteins. (I) has enhanced properties such as substantially enhanced fluorescence and reduced toxicity. This is the amino acid sequence of RedII, a mutant of the red fluorescent protein isolated from an aquatic species believed to be either an *Actinodiscus* or *Discosoma* species.

Sequence 236 AA;

Query Match 98.5%; Score 1196; DB 8; Length 236;  
 Best Local Similarity 98.2%; Pred. No. 1.2e-125;  
 Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSSKNVIVKEFMRFKVRMEGTGNGHGFELIEGEGGRPRYEGHNTVYLKTKGSPLPFANDI 60  
 DB 1 MSCSKNVIVKEFMRFKVRMEGTGNGHGFELIEGEGGRPRYEGHNTVYLKTKGSPLPFANDI 60  
 QY 61 LSPQFOYGSKYVYVGPADIPYKLSFPEGFKMERVMNFBGCVTVTVTQDSLDGCFIY 120  
 DB 61 LSPQFOYGSKYVYVGPADIPYKLSFPEGFKMERVMNFBGCVTVTVTQDSLDGCFIY 120  
 QY 121 KVFIFGVNFPSPDGPVMQKKTGMGEASTERLPRDGVLTGKEIHKALKDGGHYLVEFKSI 180  
 DB 121 KVFIFGVNFPSPDGPVMQKKTGMGEASTERLPRDGVLTGKEIHKALKDGGHYLVEFKTI 180  
 QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 DB 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 67

ADZ84223  
 ID ADZ84223 standard; protein; 236 AA.

ADZ84223;  
 DT 14-JUL-2005 (first entry)

DE Actinodiscus/Discosoma RFP mutant Red I, SEQ ID NO:2.

KW Red fluorescent protein; Red I; muteln.

OS Corallimorpharia.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 232 /note= "This residue is Asn in the mutant Ac/DsRFP Red

FT II"

XX US2005100954-A1.  
 PN 12-MAY-2005.  
 XX 23-DEC-2004; 2004US-00021014.  
 XX 09-DEC-2002; 2002US-00314936.  
 PR (GIBB//) GIBBS P D L.  
 PA (CART//) CARTER R W.  
 PA (SCHEM//) SCHMALE M C.  
 XX GIBBS PDL, Carter RW, Schmale MC;  
 DR WPI; 2005-345399/35.  
 DR N-PsDB; ADZ84222.  
 PT Novel mutant fluorescent protein comprising amino acid sequence that  
 PT differs from wild-type red or green fluorescent protein by amino acid  
 PT substitutions, useful as markers for detecting desired gene expression.  
 PS Claim 6; SEQ ID NO 2; 29pp; English.

The invention relates to mutant red and green fluorescent proteins which have higher fluorescent intensities compared to the wild-type proteins from which they are derived. The invention also discloses nucleic acids encoding the mutant fluorescent proteins, and transgenic animals comprising such a nucleic acid. The mutant red fluorescent protein (RFP) of the invention, Red I (ADZ84223) and Red II (ADZ84225), are derived from a novel wild-type RFP (referred to as Ac/DsRFP) isolated from a mushroom coral believed to be either an *Actinodiscus* or *Discosoma* species, while the mutant green fluorescent proteins (GFP), Green I (ADZ84227) and Green (ADZ84229), are derived from a novel wild-type GFP (known as MCGFP) isolated from the great star coral *Montastraea cavernosa*. The Ac/DsRFP and MCGFP mutants were obtained by subjecting the wild-type Ac/DsRFP and MCGFP cDNAs to low-stringency PCR to introduce random mutations, cloning the PCR products into bacterial expression vectors, and selecting the transformed bacteria having higher relative fluorescence compared to those expressing the wild-type proteins. The mutant fluorescent proteins of the invention may be used in a wide range of applications. They are useful as markers for detecting the expression of a gene of interest, or can be used as in vivo markers in mRNA microinjection assays in transgenic animals. They can be used in expression systems (e.g., as components of fusion proteins), in multiple labeling systems, or can be used in biochemical assays and as reagents. They may additionally be used as taxonomic markers for studies of cnidarian genetics, as color indicators in diagnostic kits, as colored food additives and as cosmetic ingredients. The present sequence represents the mutant *Actinodiscus/Discosoma* red fluorescent protein Red I.

Sequence 236 AA;

Query Match 98.5%; Score 1196; DB 9; Length 236;  
 Best Local Similarity 98.2%; Pred. No. 1.2e-125;  
 Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRSSKNVIVKEFMRFKVRMEGTGNGHGFELIEGEGGRPRYEGHNTVYLKTKGSPLPFANDI 60  
 DB 1 MSCSKNVIVKEFMRFKVRMEGTGNGHGFELIEGEGGRPRYEGHNTVYLKTKGSPLPFANDI 60  
 QY 61 LSPQFOYGSKYVYVGPADIPYKLSFPEGFKMERVMNFBGCVTVTVTQDSLDGCFIY 120  
 DB 61 LSPQFOYGSKYVYVGPADIPYKLSFPEGFKMERVMNFBGCVTVTVTQDSLDGCFIY 120  
 QY 121 KVFIFGVNFPSPDGPVMQKKTGMGEASTERLPRDGVLTGKEIHKALKDGGHYLVEFKSI 180  
 DB 121 KVFIFGVNFPSPDGPVMQKKTGMGEASTERLPRDGVLTGKEIHKALKDGGHYLVEFKTI 180  
 QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 DB 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 68  
AD284225  
ID AD284225 standard; protein; 236 AA.  
XX  
XX AD284225;  
XX  
XX 14-JUL-2005 (first entry)  
XX  
XX Actinodiscus/Discosoma RFP mutant Red II, SEQ ID NO.4.  
XX  
XX Red fluorescent protein; Red II; mtein.  
XX  
XX Corallimorpharia.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
XX FT MISC-difference 232  
XX FT /note="This residue is Asp in the mutant Ac/DsRFP Red I"  
XX  
XX PN US2005100954-A1.  
XX  
XX 12-MAY-2005.  
XX  
XX 23-DEC-2004; 2004US-00021014.  
XX  
XX 09-DEC-2002; 2002US-00314936.  
XX  
XX (GIBB/) GIBBS P D L.  
XX PA (CART/) CARTER R W.  
XX PA (SCHW/) SCHWALZ M C.  
XX  
XX GIBBS PDL, Carter RW, Schmale MC;  
XX  
XX WPI; 2005-345399/35.  
XX N-PSDB; AD284224.  
XX  
XX Novel mutant fluorescent protein comprising amino acid sequence that  
XX differs from wild-type red or green fluorescent protein by amino acid  
XX substitutions, useful as markers for detecting desired gene expression.  
XX  
XX Claim 7; SEQ ID NO 4; 29pp; English.  
XX  
XX The invention relates to mutant red and green fluorescent proteins which  
XX have higher fluorescent intensities compared to the wild-type proteins  
XX from which they are derived. The invention also discloses nucleic acids  
XX encoding the mutant fluorescent proteins, and transgenic animals  
XX comprising such a nucleic acid. The mutant red fluorescent proteins (RFP)  
XX of the invention, Red I (AD284223) and Red II (AD284225), are derived  
XX from a novel wild-type RFP (referred to as Ac/DsRFP) isolated from a  
XX mushroom coral believed to be either an Actinodiscus or Discosoma  
XX species, while the mutant green fluorescent proteins (GFP), Green I  
XX (AD284227) and Green (AD284229), are derived from a novel wild-type GFP  
XX (known as McGFP) isolated from the great star coral Montastraea  
XX cavernosa. The Ac/DsRFP and McGFP mutants were obtained by subjecting the  
XX wild-type Ac/DsRFP and McGFP cDNAs to low-stringency PCR to introduce  
XX random mutations, cloning the PCR products into bacterial expression  
XX vectors, and selecting the transformed bacteria having higher relative  
XX fluorescence compared to those expressing the wild-type proteins. The  
XX mutant fluorescent proteins of the invention may be used in a wide range  
XX of applications. They are useful as markers for detecting the expression  
XX of a gene of interest, or can be used as in vivo markers in mRNA  
XX microinjection assays in transgenic animals. They can be used in  
XX expression systems (e.g., as components of fusion proteins), in multiple  
XX labeling systems, or can be used in biochemical assays and as reagents.  
XX They may additionally be used as taxonomic markers for studies of  
XX chondarian genetics, as color indicators in diagnostic kits, as colored  
XX food additives and as cosmetic ingredients. The present sequence  
XX represents the mutant Actinodiscus/Discosoma red fluorescent protein Red  
XX II. This protein was found to have a fluorescent intensity at least 50%  
XX greater than the mutant RFP Red I (AD284223).

SEQ Sequence 236 AA;  
Query Match 98.5%; Score 1196; DB 9; Length 236;  
Best Local Similarity 98.2%; Pred. No. 1.2e-125;  
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MSSKNVTKKEPFRFVRMEGTNNGHEPEIEBGEGRPYEGHNTVTKAKTKGGPLPFPAMD 60  
DB 1 MSCSKNVTKKEPFRFVRMEGTNNGHEPEIEBGEGRPYEGHNTVTKAKTKGGPLPFPAMD 60  
QY 61 LSPFOYGSKVVYKVPADIPDYKLSFPEGFMEKRVNMFEDGGVTVYTQDSSLQDGCIFY 120  
DB 61 LSPFOYGSKVVYKVPADIPDYKLSFPEGFMEKRVNMFEDGGVTVYTQDSSLQDGCIFY 120  
QY 121 KYKFTGVNFPSPDGPVMOKKTWMEASTERLYPRDGVLCGEIKALKLKDGGHYLVEFKSI 180  
DB 121 KYKFTGVNFPSPDGPVMOKKTWMEASTERLYPRDGVLCGEIKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKRPVQLPGYVVYDSKLDITSHNEDYTVBOYERTEGRHHFL 225  
DB 181 YMAKRPVQLPGYVVYDSKLDITSHNEDYTVBOYERTEGRHHFL 225  
RESULT 69  
ADQ59562  
ID ADQ59562 standard; protein; 284 AA.  
XX  
XX ADQ59562;  
XX  
XX 07-OCT-2004 (first entry)  
XX  
XX Discosoma fluorescent protein drfp583.  
XX  
XX optically detectable signal; protein fragment complementation assay; PCA;  
XX molecular interaction detection; biomolecular interaction detection;  
XX fluorescent protein; multi-colour PCA; drug discovery; target validation;  
XX high-throughput screening; high-content screening; pathway mapping;  
XX drug mechanism-of-action study; biosensor; diagnostic;  
XX fluorescent protein; spectral shift.  
XX  
XX Discosoma.  
XX  
XX US2004137528-A1.  
XX  
XX 15-JUL-2004.  
XX  
XX 01-DEC-2003; 2003US-00724178.  
XX  
XX 02-FEB-1998; 98US-00017412.  
XX 07-FEB-2000; 2000US-00499464.  
XX 24-MAY-2002; 2002US-00154758.  
XX 29-JAN-2003; 2003US-00353090.  
XX 09-APR-2003; 2003US-0461133P.  
XX  
XX (MICH/) WATSON MICHNICK S W.  
XX PA (MACD/) MACDONALD M L.  
XX PA (LAME/) LAMERDIN J.  
XX  
XX Watson Michnick SW, Macdonald ML, Lamerdin J;  
XX  
XX WPI; 2004-533363/51.  
XX  
XX Composition useful in protein fragment complementation assays for drug  
XX discovery and high-throughput screening, comprising complementary  
XX fragments of protein or mutant protein, generating optically detectable  
XX signal when associated.  
XX  
XX Disclosure; SEQ ID NO 12; 34pp; English.  
XX  
XX The invention describes a composition (I) comprising complementary  
XX fragments of a protein or mutant protein, where the fragments generate an  
XX optically detectable signal when associated, and each of the mutant  
XX protein fragments is fused to a separate molecule. Also described are:

CC protein fragment complementation assays (PCAs) for detection of molecular  
 CC interactions, involving reassembling separate fragments from an optically  
 CC detectable protein, and detecting the reassembly by units of  
 CC reconstruction of activity of the optically detectable protein, where the  
 CC reassembly of the fragments is operated by the interaction of molecular  
 CC domains fused to each fragment, and is independent of other molecular  
 CC processes; detecting biomolecular interaction, involving selecting an  
 CC appropriate optically detectable protein, effecting fragmentation of the  
 CC optically detectable protein such that the fragmentation results in  
 CC reversible loss of protein function, fusing or attaching fragments of the  
 CC optically detectable protein separately to other molecules, reassociating  
 CC the protein fragments through interactions of the molecules that are  
 CC fused or attached to the fragments, and detecting the resulting optical  
 CC signal; designing and engineering of PCAs based on fluorescent protein;  
 CC and a method and composition for the construction of multi-color PCAs.  
 CC (1) is useful in PCAs and other assays for drug discovery, target  
 CC validation, high-throughput screening, high-content screening, pathway  
 CC mapping, drug mechanism-of-action studies, biosensors and diagnostics.  
 CC (1) is useful for engineering different colour PCAs for a variety of  
 CC applications in biology and biotechnology. This is the amino acid  
 CC sequence of a Discosoma fluorescent protein used in a comparison of  
 CC fluorescent proteins with Aequorea green fluorescent protein, variants of  
 CC which can be created with altered excitation and emission wavelength  
 CC relative to wild type GFP.

CC Sequence 284 AA;

Query Match 98.5%; Score 1195.5; DB 8; Length 284;

Best Local Similarity 99.1%; Pred. No. 1.8e-125; Indels 1; Gaps 1;

Matches 224; Conservative 0; Mismatches 1;

QY 1 MRSSKNVKEPFRFKVMEGTNGHGFIEGEGGRPYGHNTYKLTQTKGSPLEPFANDI 60  
 DB 59 MRSSKNVKEPFRFKVMEGTNGHGFIEGEGGRPYGHNTYKLTQTKGSPLEPFANDI 118  
 QY 61 LSPQFQYSGSKYVVKRPADIPDYKKLSFPEGFKMERVWVNFEDGGVTVTQDSSLQDGCFTY 120  
 DB 119 LSPQFQYSGSKYVVKRPADIPDYKKLSFPEGFKMERVWVNFEDGGVTVTQDSSLQDGCFTY 178  
 QY 121 KYKFLGVNFPSPDGPVWQKKTWGMWASTERTLYPRDGLKGEIHKALKLDGGHYLVEFQSI 180  
 DB 179 KYKFLGVNFPSPDGPVWQKKTWGMWASTERTLYPRDGLKGEIHKALKLDGGHYLVEFQSI 238  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
 DB 239 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 284

RESULT 70

ABP56681  
 ID ABP56681 standard; protein; 240 AA.

AC ABP56681;

XX 25-MAR-2003 (first entry)

DE Discosoma red fluorescent protein variant F125L and N24S.

KM Mammalian codon optimised Discosoma red fluorescent protein; variant;

XX Discosoma; red fluorescent protein; directed evolution; mutant.

OS Discosoma sp.

XX Synthetic.

XX WO200294992-A2.

XX 28-NOV-2002.

XX 20-MAY-2002; 2002WO-US015968.

XX 18-MAY-2001; 2001US-0291871P.

XX (RIGI-) RIGEL PHARM INC.

XX Peelle B;  
 PI  
 XX WPI; 2003-120798/11.

PT New Discosoma red fluorescent protein, useful for functional screens as a  
 PT reporter for gene transcription, for target characterization and  
 PT localization of fusion proteins, or for scaffolds for protein and peptide  
 PT libraries.

PS Claim 11; Page; 22pp; English.

CC The present invention describes an isolated Discosoma red fluorescent  
 CC protein (1) comprising a 241 residue amino acid sequence (see ABP56678,  
 CC S1), with one or more point mutations at amino acid position N24, F125,  
 CC K164, or M183. Also described: (1) a fusion protein comprising (1); (2)  
 CC an isolated nucleic acid encoding (1); (3) a vector comprising the  
 CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
 CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
 CC encoding (1); and (6) methods of making a fluorescent variant. The  
 CC Discosoma red fluorescent proteins are useful for functional screens as a  
 CC reporter for gene transcription (e.g. as a fusion protein), for target  
 CC characterisation and localisation of fusion proteins, or for scaffolds  
 CC for protein and peptide libraries. The fluorescent proteins can also be  
 CC used as selectable markers or reporter molecules for a variety of  
 CC bioassays, including methods that use fluorescence activated cell sorting  
 CC (FACS) as a selection mechanism. The method of directed protein evolution  
 CC is useful for obtaining improved variants of red fluorescent protein. The  
 CC variants of Discosoma red fluorescent protein have greatly improved  
 CC brightness, expression, and/or folding kinetics as compared to wild type  
 CC or a codon optimised variant. The present sequence represents a mammalian  
 CC codon optimised Discosoma red fluorescent protein variant from the  
 CC present invention. N.B. The present sequence is not given in the  
 CC specification, but is derived from the protein in Fig 1 as stated in the  
 CC claims

CC Sequence 240 AA;

Query Match 98.4%; Score 1194; DB 6; Length 240;

Best Local Similarity 98.2%; Pred. No. 2e-125; Indels 0; Gaps 0;

Matches 221; Conservative 3; Mismatches 1;

QY 1 MRSSKNVKEPFRFKVMEGTNGHGFIEGEGGRPYGHNTYKLTQTKGSPLEPFANDI 60  
 DB 2 VRSSKNVKEPFRFKVMEGTNGHGFIEGEGGRPYGHNTYKLTQTKGSPLEPFANDI 61  
 QY 61 LSPQFQYSGSKYVVKRPADIPDYKKLSFPEGFKMERVWVNFEDGGVTVTQDSSLQDGCFTY 120  
 DB 62 LSPQFQYSGSKYVVKRPADIPDYKKLSFPEGFKMERVWVNFEDGGVTVTQDSSLQDGCFTY 121  
 QY 121 KYKFLGVNFPSPDGPVWQKKTWGMWASTERTLYPRDGLKGEIHKALKLDGGHYLVEFQSI 180  
 DB 122 KYKFLGVNFPSPDGPVWQKKTWGMWASTERTLYPRDGLKGEIHKALKLDGGHYLVEFQSI 181  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
 DB 182 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 226

RESULT 71

ADH34506  
 ID ADH34506 standard; protein; 225 AA.

AC ADH34506;

XX 11-MAR-2004 (first entry)

DE Discosoma sp. Dated mutant Dated-NF S148C/I165N/K167M/S203A.

KM Chromoprotein, fluorescent protein; CP; FP; interconverted mutant;

KM Chidnarian; Anthozoan; labeling; colouring agents; pigment;

KM analyte detection assay; selectable marker; sunscreen; selective filter;

KM fluorescence resonance energy transfer; FRRET; biosensor;



KW whole cell marker; second messenger detector; in vivo marker;  
 KW fluorescence activated cell sorting; fluorescent timer;  
 KW red fluorescent protein; Dared; non-fluorescent mutant; Dared-NF; mutant;  
 KW muten.  
 XX Synthetic.  
 OS Discosoma sp.  
 FH Key  
 FT Location/Qualifiers  
 FT M18c-difference 146  
 FT /note= "lys replaces w1d-type Ser. Corresponds to GFP  
 FT residue 148"  
 FT M18c-difference 161  
 FT /note= "Asn replaces w1d-type Ile. Corresponds to GFP  
 FT residue 165"  
 FT M18c-difference 163  
 FT /note= "Met replaces w1d-type Lys. Corresponds to GFP  
 FT residue 167"  
 FT M18c-difference 197  
 FT /note= "Ala replaces w1d-type Ser. Corresponds to GFP  
 FT residue 203"  
 XX MO2003057833-A2.  
 XX 17-JUL-2003.  
 XX 23-DEC-2002; 2002MO-US041418.  
 XX 26-DEC-2001; 2001US-0343128P.  
 (CLON-) CLONTECH LAB INC.  
 PA Bulina ME, Chudakov D, Lukyanov KA;  
 PI MPI; 2003-607998/57.  
 DR Novel nucleic acid encoding interconverted mutant of chromo-or  
 PT fluorescent protein which are useful as biosensors, coloring agents.  
 XX Example 1; Page: 56pp; English.  
 PS The invention relates to interconverted mutants of chromoproteins (CP) or  
 CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
 CC derived from a Cnidarian species, preferably a non-bioluminescent  
 CC Cnidarian species, and most preferably an Anthozoan species. The  
 CC invention is based on the finding that although green fluorescent protein  
 CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
 CC homology, there are certain positions (referred to as 148, 165, 167 and  
 CC 203; numbering corresponds to GFP) that are occupied by noticeably  
 CC different residues in the two types of proteins. Mutagenesis of the  
 CC residues in these key positions in, for example, a fluorescent protein,  
 CC to those found in a chromoprotein is therefore proposed to confer  
 CC chromoprotein activity on the fluorescent protein mutant, with  
 CC chromoproteins being able to be converted into fluorescent proteins in a  
 CC similar manner. The invention also relates to expression constructs,  
 CC vectors, host cells and host cell progeny comprising a nucleic acid of  
 CC the invention; the recombinant production of an interconverted  
 CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
 CC interconverted mutant proteins of the invention. The interconverted  
 CC mutants are useful in any application that employs a chromoprotein or  
 CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
 CC activity can be useful as coloring agents in, for example, food  
 CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
 CC with chromoprotein activity are also useful as labels in biological  
 CC analysis detection assays, as selectable markers in recombinant DNA  
 CC applications (e.g. the production of transgenic cells and organisms), and  
 CC are also useful as biosensors and selective filters. Chromoproteins  
 CC mutants having fluorescent protein activity useful in fluorescence  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
 CC changes in multicellular reorganisation and migration, as second  
 CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
 CC animals), in fluorescence activated cell sorting applications, in

CC protease cleavage assays, and in assays to determine the phospholipid  
 CC composition in biological membranes. Proteins with fluorescent protein  
 CC activity can also be used as fluorescent timers, where the switch of one  
 CC fluorescent colour to another (e.g., green to red) is concomitant with  
 CC the ageing of the protein and is useful for determination of the  
 CC activation or deactivation of gene expression. The present sequence  
 CC represents a non-fluorescent Discosoma sp. red fluorescent protein Dared  
 CC mutant, Dared-NF, which has chromoprotein activity and which was  
 CC generated in an example of the invention. The present sequence is not  
 CC shown in the specification, but was derived from the wild-type Dared  
 CC sequence (ADH34489) shown in Fig 1 and the information provided on page  
 CC 42.  
 XX SQ Sequence 225 AA;  
 XX  
 XX Query Match 98.3%; Score 1193; DB 7; Length 225;  
 XX Best Local Similarity 98.2%; Pred. No. 2,4e-125;  
 XX Matches 221; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MRSKNVKEFMRFKVRMEGTNGHFEIEGEGEGRPYGHNTVTLKXTKGGPLPAPMDI 60  
 DB 1 MRSKNVKEFMRFKVRMEGTNGHFEIEGEGEGRPYGHNTVTLKXTKGGPLPAPMDI 60  
 QY 61 LSPPOYGSKYVVKIPADIPDYKXLSFPEGFKMERVMNFDGAVTVTQDSSLQDGCFTY 120  
 DB 61 LSPPOYGSKYVVKIPADIPDYKXLSFPEGFKMERVMNFDGAVTVTQDSSLQDGCFTY 120  
 QY 121 KYKFIGVNPSPDGPVWQKKTWGEASTERYLPDGVLKGEIHKALKDKGHIYVEFKSI 180  
 DB 121 KYKFIGVNPSPDGPVWQKKTWGEASTERYLPDGVLKGEIHKALKDKGHIYVEFKSI 180  
 QY 181 YMAKKPVQLPGYIYDSDKDITSNEDTYIVQYERTBERHHLFL 225  
 DB 181 YMAKKPVQLPGYIYDSDKDITSNEDTYIVQYERTBERHHLFL 225  
 DB 181 YMAKKPVQLPGYIYDSDKDITSNEDTYIVQYERTBERHHLFL 225  
 RESULT 72  
 ABP56679  
 ID ABP56679 standard; protein; 240 AA.  
 XX  
 AC ABP56679;  
 XX  
 DT 25-MAR-2003 (first entry)  
 XX  
 DE Discosoma red fluorescent protein variant P125L and M183K.  
 XX  
 KW Mammalian codon optimised Discosoma red fluorescent protein; variant;  
 KW Discosoma; red fluorescent protein; directed evolution; mutant.  
 XX  
 OS Discosoma sp.  
 OS Synthetic.  
 XX  
 PN MO200294992-A2.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 20-MAY-2002; 2002MO-US015968.  
 XX  
 PR 18-MAY-2001; 2001US-0291871P.  
 XX  
 PA (RIGL-) RIGEL PHARM INC.  
 PI Peelle B;  
 XX  
 DR MPI; 2003-120798/11.  
 XX  
 PT New Discosoma red fluorescent protein, useful for functional screens as a  
 PT reporter for gene transcription, for target characterization and  
 PT localization of fusion proteins, or for scaffolds for protein and peptide  
 PT libraries.  
 XX  
 XX Claim 8; Page: 22pp; English.  
 XX

CC The present invention describes an isolated *Discosoma* red fluorescent  
CC protein (1) comprising a 241 residue amino acid sequence (see ABP56678,  
CC S1), with one or more point mutations at amino acid position N24, F125,  
CC K164, or M183. Also described: (1) a fusion protein comprising (1); (2)  
CC an isolated nucleic acid encoding (1); (3) a vector comprising the  
CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
CC encoding (1); and (6) methods of making a fluorescent variant. The  
CC *Discosoma* red fluorescent proteins are useful for functional screens as a  
CC reporter for gene transcription (e.g. as a fusion protein), for target  
CC characterisation and localisation of fusion proteins, or for scaffolds  
CC for protein and peptide libraries. The fluorescent proteins can also be  
CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of *Discosoma* red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised *Discosoma* red fluorescent protein variant from the  
CC present invention. N.B. The present sequence is not given in the  
CC specification, but is derived from the protein in Fig 1 as stated in the  
CC claims

CC Sequence 240 AA;

Query Match 98.3%; Score 1193; DB 6; Length 240;  
Best Local Similarity 98.2%; Pred. No. 2,6e-125;  
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRFRKVRMEGTNGHBEFIEGEGGRPYGHNTVYKLYTKGGLPFANDI 60  
DB 2 VRSSKNVKEFRFRKVRMEGTNGHBEFIEGEGGRPYGHNTVYKLYTKGGLPFANDI 61  
QY 61 LSPQFQYGSKYVYVGRPADIPDYKKLSPEEGFKWRVWNNFEDGCVTVYQDSSLQDGCFTY 120  
DB 62 LSPQFQYGSKYVYVGRPADIPDYKKLSPEEGFKWRVWNNFEDGCVTVYQDSSLQDGCFTY 121  
QY 121 KVKFIGNVPSPDGPVWQKKTWGMESTERLYPRDGLKGEIHKALKLDGSHYLVFPSI 180  
DB 122 KVKFIGNVPSPDGPVWQKKTWGMESTERLYPRDGLKGEIHKALKLDGSHYLVFPSI 181  
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTYIVEQYERTEGRHHLFL 225  
DB 182 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTYIVEQYERTEGRHHLFL 226

RESULT 73

ABP56680  
ID ABP56680 standard; protein; 240 AA.

AC ABP56680;

DT 25-MAR-2003 (first entry)

DE *Discosoma* red fluorescent protein variant F125V and M183K.

KM Mammalian codon optimised *Discosoma* red fluorescent protein; variant;  
KW *Discosoma*; red fluorescent protein; directed evolution; mutant.

OS *Discosoma* sp.  
OS Synthetic.

FN WO200294992-A2.

PD 28-NOV-2002.

PF 20-MAY-2002; 2002WO-US015968.

PR 18-MAY-2001; 2001US-0291871P.

PA (RIGE-) RIGEL PHARM INC.

XX

PI Peeble B;  
XX WPI; 2003-120798/11.  
DR  
XX  
PT New *Discosoma* red fluorescent protein, useful for functional screens as a  
PT reporter for gene transcription, for target characterization and  
PT localisation of fusion proteins, or for scaffolds for protein and peptide  
PT libraries.

PS Claim 9; Page: 22pp; English.

CC The present invention describes an isolated *Discosoma* red fluorescent  
CC protein (1) comprising a 241 residue amino acid sequence (see ABP56678,  
CC S1), with one or more point mutations at amino acid position N24, F125,  
CC K164, or M183. Also described: (1) a fusion protein comprising (1); (2)  
CC an isolated nucleic acid encoding (1); (3) a vector comprising the  
CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
CC encoding (1); and (6) methods of making a fluorescent variant. The  
CC *Discosoma* red fluorescent proteins are useful for functional screens as a  
CC reporter for gene transcription (e.g. as a fusion protein), for target  
CC characterisation and localisation of fusion proteins, or for scaffolds  
CC for protein and peptide libraries. The fluorescent proteins can also be  
CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of *Discosoma* red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised *Discosoma* red fluorescent protein variant from the  
CC present invention. N.B. The present sequence is not given in the  
CC specification, but is derived from the protein in Fig 1 as stated in the  
CC claims

CC Sequence 240 AA;

Query Match 98.2%; Score 1192; DB 6; Length 240;  
Best Local Similarity 98.2%; Pred. No. 3,4e-125;  
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRFRKVRMEGTNGHBEFIEGEGGRPYGHNTVYKLYTKGGLPFANDI 60  
DB 2 VRSSKNVKEFRFRKVRMEGTNGHBEFIEGEGGRPYGHNTVYKLYTKGGLPFANDI 61  
QY 61 LSPQFQYGSKYVYVGRPADIPDYKKLSPEEGFKWRVWNNFEDGCVTVYQDSSLQDGCFTY 120  
DB 62 LSPQFQYGSKYVYVGRPADIPDYKKLSPEEGFKWRVWNNFEDGCVTVYQDSSLQDGCFTY 121  
QY 121 KVKFIGNVPSPDGPVWQKKTWGMESTERLYPRDGLKGEIHKALKLDGSHYLVFPSI 180  
DB 122 KVKFIGNVPSPDGPVWQKKTWGMESTERLYPRDGLKGEIHKALKLDGSHYLVFPSI 181  
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTYIVEQYERTEGRHHLFL 225  
DB 182 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTYIVEQYERTEGRHHLFL 226

RESULT 74

AAE28837  
ID AAE28837 standard; protein; 225 AA.

AC AAE28837;

DT 27-DEC-2002 (first entry)

DE *Discosoma* sp. drFP583 (NFP-6) mutant protein, E5-NA.

KM Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRRT;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;

XX	mutant.
XX	
OS	Discooma sp.
OS	Synthetic.
XX	
XX	Key
FT	Misc-difference 2
FT	Location/Qualifiers
FT	/note= "Wild type Arg substituted with Ala"
FT	Misc-difference 5
FT	/note= "Wild type Lys substituted with Glu"
FT	Misc-difference 9
FT	/note= "Wild type Lys substituted with Thr"
FT	Misc-difference 105
FT	/note= "Wild type Val substituted with Ala"
FT	Misc-difference 197
FT	/note= "Wild type Ser substituted with Thr"
XX	
PN	WO200268459-A2.
PD	
XX	06-SEP-2002.
PR	20-FEB-2004; 2002MO-US005749.
XX	
XX	21-FEB-2001; 2001US-0270983P.
PR	04-DEC-2001; 2001US-00006922.
XX	
PA	(CLON-) CLONTECH LAB INC.
PI	
XX	Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Pradkov A;
DR	WPI; 2002-691654/74.
XX	
DR	N-PSDB; AAD46282.
XX	
PT	New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant
PT	of an aggregating Crdarian chromo- or fluorescent protein or mutant for
PT	analyte detection assays or fluorescence activated cell sorting
XX	applications.
XX	
PS	Disclosure; Page; 80pp; English.
XX	
CC	The invention relates to nucleic acid molecules encoding non-aggregating
CC	chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are
CC	useful in analyte detection assays, as colouring agents, as markers in
CC	recombinant DNA applications, as screens or filters, in fluorescence
CC	resonance energy transfer (FRET) applications, as biosensors in
CC	prokaryotic and eukaryotic cells, in screening assays, as second
CC	messenger detectors, in fluorescence activated cell sorting applications,
CC	in protease cleavage assays or as fluorescent timers. The present
CC	invention. Note: This sequence is not shown in the specification, but is
CC	derived from Discooma sp. dirp583 (NFP-6) wild-type protein shown as SEQ
CC	ID NO:8 (AAE28833) in page 70-71 of the specification
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	
XX	

Db 181 YMAKPVQLEPGYYVDTKIDITSNEDYTVIEQYERTEGRHHFL 225

RESULT 75

ID AAE17542 standard; protein; 225 AA.

XX AAE17542;

AC AAE17542;

XX AAE17542;

DT 22-APR-2002 (first entry)

DE Discoosoma sp. non-aggregating mutant timer protein ESNA.

XX Fluorescent timer protein; protein movement; translocation; trafficking;

KW promoter activity; gene expression; transgenic plant; gene modification;

KW protein age; mutant; ESNA; muten.

XX

OS Discoosoma sp.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 2 /note= "Wild type Arg substituted with Ala"

FT Misc-difference 5 /note= "Wild type Lys substituted with Glu"

FT Misc-difference 9 /note= "Wild type Lys substituted with Thr"

FT

PN MO200196373-A2.

PD 20-DEC-2001.

XX

XX 13-JUN-2001; 2001MO-US019097.

PF 14-JUN-2000; 2000US-0211607P.

PR

XX (CLON-) CLONTECH LAB INC.

PA

PI Fradkov AF, Terakikh A;

XX

DR MPI: 2002-154595/20.

XX N-PSDB; MAD28209.

XX

PT New fluorescent timer proteins comprising an emission spectrum that

PT changes over time from a first wavelength to a second wavelength, useful

PT for monitoring intracellular protein movement, translocation, trafficking

PT or stability.

XX

PS Claim 5, Fig 3, 89pp; English.

XX

XX The invention relates to a fluorescent timer protein having an emission

XX spectrum that changes over time after synthesis from a first wavelength

XX to a second wavelength. The fluorescent timer proteins are useful in

XX monitoring the activity of a promoter, determining the age of a protein,

XX identifying an agent that modulates the activity of a promoter and in

XX enriching a population of cells comprising a fluorescent timer protein.

XX The fluorescent timer proteins are also useful for assessing gene

XX expression during development of a multicellular organism or during

XX cellular differentiation, in response to a drug or other inducer of

XX promoter activity, as a reporter to serve as a read-out of promoter

XX activity, monitoring intracellular protein movement or translocation,

XX protein trafficking, or protein stability, to investigate temporal

XX aspects of the activity of a regulatory element, for determining cell

XX fate during development and organ remodelling, in spatial and temporal

XX visualisation of newly synthesised proteins and accumulated proteins, and

XX in distinguishing between newly formed and pre-existing structures, e.g.

XX membrane junctions and extracellular matrix components. The fluorescent

XX timer proteins may further be used to investigate systems where photobleaching

XX techniques are employed, as detectable labels, as selectable markers, as

XX biosensors in prokaryotic and eukaryotic cells, in protease cleavage

XX assays, and as second messenger detectors. The nucleic acids can be used

XX to generate transgenic, non-human plants or animals or site-specific gene

XX modifications in cell lines. The present sequence is Discoosoma sp. ESNA

CC non-aggregating fluorescent timer protein derived from fluorescent timer  
 CC protein E5 by substituting amino acids at positions R2A, K5B and K9T. E5  
 CC protein is derived from humanised wild-type Anthozoa protein drFP583 by  
 CC substituting Val to Ala at 105 and Ser to Thr at 197

XX Sequence 225 AA;

Query Match 98.1%; Score 1191; DB 5; Length 225;  
 Best Local Similarity 97.8%; Pred. No. 4e-125;  
 Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRSSKNVYKEFMRPKVMEGTVNGHEFEIEGEGRPYEGHNTVYKLTQYKGGPLPFAMDI 60  
 1 MASSENVITEFMRPKVMEGTVNGHEFEIEGEGRPYEGHNTVYKLTQYKGGPLPFAMDI 60  
 DB 61 LSPQFQYGSKYVYVGRPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 120  
 61 LSPQFQYGSKYVYVGRPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 120  
 QY 121 KYKFTGVNFPDSGPMQKTMGMEASTERLYPRDGVLKGEIHKALKDGGHYLVEFQSI 180  
 121 KYKFTGVNFPDSGPMQKTMGMEASTERLYPRDGVLKGEIHKALKDGGHYLVEFQSI 180  
 DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

#### RESULT 76

ABP56682  
 ID ABP56682 standard; protein; 240 AA.

XX ABP56682;

DT 25-MAR-2003 (first entry)

XX Discosoma red fluorescent protein variant F125L, N24S and M183K.

XX Mammalian codon optimised Discosoma red fluorescent protein; variant;

KM Discosoma; red fluorescent protein; directed evolution; mutant.

XX Discosoma sp.

OS Synthetic.

XX WO200294992-A2.

PN 28-NOV-2002.

PD 20-MAY-2002; 2002WO-US015968.

PF 18-MAY-2001; 2001US-0291871P.

PR (RIGE-) RIGEL PHARM INC.

XX Peelie B;

PI WPI; 2003-120798/11.

DR New Discosoma red fluorescent protein, useful for functional screens as a  
 PT reporter for gene transcription, for target characterization and  
 PT localization of fusion proteins, or for scaffolds for protein and peptide  
 PT libraries.

XX Claim 13; Page: 22pp; English.

XX The present invention describes an isolated Discosoma red fluorescent  
 CC protein (1) comprising a 241 residue amino acid sequence (see ABP56678,  
 CC S1) with one or more point mutations at amino acid position N24, F125,  
 CC K164, or M183. Also described: (1) a fusion protein comprising (1); (2)  
 CC an isolated nucleic acid encoding (1); (3) a vector comprising the  
 CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
 CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
 CC encoding (1); and (6) methods of making a fluorescent variant. The

CC Discosoma red fluorescent proteins are useful for functional screens as a  
 CC reporter for gene transcription (e.g. as a fusion protein), for target  
 CC characterisation and localisation of fusion proteins, or for scaffolds  
 CC for protein and peptide libraries. The fluorescent proteins can also be  
 CC used as selectable markers or reporter molecules for a variety of  
 CC bioassays, including methods that use fluorescence activated cell sorting  
 CC (FACS) as a selection mechanism. The method of directed protein evolution  
 CC is useful for obtaining improved variants of red fluorescent protein. The  
 CC variants of Discosoma red fluorescent protein have greatly improved  
 CC brightness, expression, and/or folding kinetics as compared to wild type  
 CC or a codon optimised variant. The present sequence represents a mammalian  
 CC codon optimised Discosoma red fluorescent protein variant from the  
 CC present invention. N.B. The present sequence is not given in the  
 CC specification, but is derived from the protein in Fig 1 as stated in the  
 CC claims

XX Sequence 240 AA;

Query Match 97.9%; Score 1188; DB 6; Length 240;  
 Best Local Similarity 97.8%; Pred. No. 9.6e-125;  
 Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRSSKNVYKEFMRPKVMEGTVNGHEFEIEGEGRPYEGHNTVYKLTQYKGGPLPFAMDI 60  
 2 VRSSKNVYKEFMRPKVMEGTVNGHEFEIEGEGRPYEGHNTVYKLTQYKGGPLPFAMDI 61  
 DB 61 LSPQFQYGSKYVYVGRPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 120  
 62 LSPQFQYGSKYVYVGRPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 121  
 QY 121 KYKFTGVNFPDSGPMQKTMGMEASTERLYPRDGVLKGEIHKALKDGGHYLVEFQSI 180  
 122 KYKFTGVNFPDSGPMQKTMGMEASTERLYPRDGVLKGEIHKALKDGGHYLVEFQSI 181  
 DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
 182 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 226  
 QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

#### RESULT 77

AAE28836  
 ID AAE28836 standard; protein; 225 AA.

XX AAE28836;

DT 27-DEC-2002 (first entry)

DE Discosoma sp. drFP583 (NFP-6) mutant protein, E57-NA.

XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
 KM fluorescence activated cell sorting application; fluorescent timer;  
 KM biosensor; fluorescence resonance energy transfer application; FRRET;  
 KM colouring agent; recombinant DNA application; analyte detection assay;  
 KM sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
 XX mutant.

XX Discosoma sp.

OS Synthetic.

XX Key

FM location/Qualifiers

FT Misc-difference 2 /note= "Wild type Arg substituted with Ala"

FT Misc-difference 5 /note= "Wild type Lys substituted with Glu"

FT Misc-difference 9 /note= "Wild type Lys substituted with Thr"

FT Misc-difference 105 /note= "Wild type Val substituted with Ala"

FT Misc-difference 161 /note= "Wild type Ile substituted with Thr"

FT Misc-difference 197 /note= "Wild type Ser substituted with Ala"

PN WO200268459-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 20-FEB-2002; 2002WO-US005749.  
 XX  
 PR 21-FEB-2001; 2001US-0270983P.  
 XX 04-DEC-2001; 2001US-00006922.  
 XX  
 PA (CLON-) CLONTECH LAB INC.  
 PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
 DR MPI; 2002-691654/74.  
 N-PSDB; AAD46281.  
 PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
 PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
 PT analyte detection assays or fluorescence activated cell sorting  
 PT applications.  
 PS Disclosure; Page; 80pp; English.  
 XX  
 CC The invention relates to nucleic acid molecules encoding non-aggregating  
 CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
 CC useful in analyte detection assays, as colouring agents, as markers in  
 CC recombinant DNA applications, as screens or filters, in fluorescence  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, in screening assays, as second  
 CC messenger detectors, in fluorescence activated cell sorting applications,  
 CC in protease cleavage assays or as fluorescent timers. The present  
 CC sequence is *Discosoma* sp. dFP583 (NFP-6) mutant protein of the  
 CC invention. Note: This sequence is not shown in the specification, but is  
 CC derived from *Discosoma* sp. dFP583 (NFP-6) wild-type protein shown as SEQ  
 CC ID NO:8 (AAB28833) in page 70-71 of the specification  
 XX  
 SQ Sequence 225 AA;  
 XX  
 Query Match 97.7%; Score 1186; DB 5; Length 225;  
 Best Local Similarity 97.3%; Pred. No. 1.5e-124;  
 Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MRSKRVVKEFMFKVYRMGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60  
 DB 1 MASSEVITTEFMFKVYRMGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60  
 QY 61 LSPQFQYGSKVYVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVVTVTODSSLODGCFTY 120  
 DB 61 LSPQFQYGSKVYVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVVTVTODSSLODGCFTY 120  
 QY 121 KYVFIGNPSPDGPVWOKTGMGEASTERLYPRDGVLKGEIHKALKKDGHYLVEFKSI 180  
 DB 121 KYVFIGNPSPDGPVWOKTGMGEASTERLYPRDGVLKGEIHKALKKDGHYLVEFKSI 180  
 QY 181 YMAKKEVQLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHHLFL 225  
 DB 181 YMAKKEVQLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHHLFL 225  
 XX  
 RESULT 78  
 ADH10227  
 ID ADH10227 standard; protein; 225 AA.  
 XX  
 AC ADH10227;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Red fluorescent protein variant (Dared).  
 XX  
 KM Protein immobilisation; red fluorescent protein; Dared.  
 XX  
 OS Synthetic.  
 XX

PN EPI367121-A2.  
 XX  
 PD 03-DEC-2003.  
 XX  
 PF 23-MAY-2003; 2003EP-00253252.  
 XX  
 PR 23-MAY-2002; 2002JP-00148950.  
 XX  
 PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
 PI Iwakura M, Hirota K;  
 DR MPI; 2004-045224/05.  
 XX  
 CC New carrier, useful for immobilizing polypeptides via the carboxyl group  
 CC of the C-terminus, comprises a protein immobilized adsorptively, and  
 CC having a sulphydryl group.  
 PS Example; SEQ ID NO 7; 31pp; English.  
 XX  
 CC The invention relates to a carrier comprising a protein immobilized  
 CC adsorptively, where the protein has a sulphydryl group. The carrier is  
 CC useful for immobilizing polypeptides via the carboxyl group of the C-  
 CC terminus. The present sequence represents a Red fluorescent protein  
 CC variant (Dared) sequence.  
 XX  
 SQ Sequence 225 AA;  
 XX  
 Query Match 97.7%; Score 1186; DB 8; Length 225;  
 Best Local Similarity 97.3%; Pred. No. 1.5e-124;  
 Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MRSKRVVKEFMFKVYRMGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60  
 DB 1 MASSEVITTEFMFKVYRMGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFPAMD 60  
 QY 61 LSPQFQYGSKVYVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVVTVTODSSLODGCFTY 120  
 DB 61 LSPQFQYGSKVYVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVVTVTODSSLODGCFTY 120  
 QY 121 KYVFIGNPSPDGPVWOKTGMGEASTERLYPRDGVLKGEIHKALKKDGHYLVEFKSI 180  
 DB 121 KYVFIGNPSPDGPVWOKTGMGEASTERLYPRDGVLKGEIHKALKKDGHYLVEFKSI 180  
 QY 181 YMAKKEVQLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHHLFL 225  
 DB 181 YMAKKEVQLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHHLFL 225  
 XX  
 RESULT 79  
 ADT77526  
 ID ADT77526 standard; protein; 225 AA.  
 XX  
 AC ADT77526;  
 XX  
 DT 13-JAN-2005 (first entry)  
 XX  
 DE Protein SEQ ID NO:2 used in protein array.  
 XX  
 KM Protein array; planar substrate; immobilisation; detection;  
 XX target substance.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2004090542-A1.  
 XX  
 PD 21-OCT-2004.  
 XX  
 PF 09-APR-2004; 2004WO-JP005150.  
 XX  
 PR 10-APR-2003; 2003JP-00106450.  
 XX  
 PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
 XX

```
XX
PI Iwakura M, Hirota K;
XX
XX WPI; 2004-766496/75.
DR
XX Substrate for protein array, obtained by combining a polymer compound
XX having primary amino groups in repeating units on surface of planar
XX substrate.
PS Example 1; SEQ ID NO 2; 32pp; Japanese.
XX
XX The invention relates to a substrate for a protein array, obtained by
XX combining a polymer compound having primary amino groups in repeating
XX units on a surface of a planar substrate. The invention also relates to a
XX protein array formed by bonding the C-termini of proteins to the primary
XX amino groups of the polymer compound on the substrate. The invention
XX further relates to a method for the preparation of the protein array.
XX Preferably, the planar substrate has water absorptivity, the polymer used
XX is a polyallyl amine or a polylysine, and the immobilised proteins
XX contain a linker peptide. The protein array of the invention enables the
XX effective and highly sensitive detection of various target substances.
XX Sequences ADT77525-ADT77526 represent proteins whose C-termini were
XX joined to a negatively charged acidic linker peptide (ADT77529) to
XX produce the sequences shown in ADT77527-ADT77528. Sequences ADT77527-
XX ADT77528 were subsequently immobilised on a planar substrate via the
XX linker peptide to form a protein array.
SQ Sequence 225 AA;

Query Match          97.7%; Score 1186; DB 8; Length 225;
Best Local Similarity 97.3%; Pred. No. 1.5e-124;
Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRSSKNVIEKEMRKFKVMEGTVNGHEFEIEGEGRPYEGHNTVYKLTXTKGSPLEPFAMDI 60
   |||:|||||
DB 1 MASSENVITEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLTXTKGSPLEPFAMDI 60

QY 61 LSPQFQYSGSKYVYVGPADIPDYKKLSFPEGFKMERVMNFEDEGVVTVTQDSSLQDGCFTY 120
   |||:|||||
DB 61 LSPQFQYSGSKYVYVGPADIPDYKKLSFPEGFKMERVMNFEDEGVVTVTQDSSLQDGCFTY 120

QY 121 KVFIFGVNFPSPDGPVWQKTMGMEASTERYLPKRGVLTGKEIHKALKLKDGGHYLVFESKI 180
   |||:|||||
DB 121 KVFIFGVNFPSPDGPVWQKTMGMEASTERYLPKRGVLTGKEIHKALKLKDGGHYLVFESKI 180

QY 121 KVFIFGVNFPSPDGPVWQKTMGMEASTERYLPKRGVLTGKEIHKALKLKDGGHYLVFESKI 180
   |||:|||||
DB 121 KVFIFGVNFPSPDGPVWQKTMGMEASTERYLPKRGVLTGKEIHKALKLKDGGHYLVFESKI 180

QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   |||:|||||
DB 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 80
ADH10228
ID ADH10228 standard; protein; 239 AA.
XX
AC ADH10228;
XX
DT 11-MAR-2004 (first entry)
XX
DE Immobilised red fluorescent protein variant (Dared).
XX
KW Protein immobilisation; red fluorescent protein; Dared.
XX
OS Synthetic.
XX
XX EP167121-A2.
XX
XX 03-DEC-2003.
XX
XX 23-MAY-2003; 2003EP-00253252.
XX
XX 23-MAY-2002; 2002JP-00148950.
XX
XX (NABD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
```

```
XX
PI Iwakura M, Hirota K;
XX
XX WPI; 2004-045224/05.
DR
XX New carrier, useful for immobilizing polypeptides via the carboxyl group
XX of the C-terminus, comprises a protein immobilized adsorptively, and
XX having a sulhydryl group.
PS Example 3; SEQ ID NO 8; 31pp; English.
XX
XX The invention relates to a carrier comprising a protein immobilized
XX adsorptively, where the protein has a sulhydryl group. The carrier is
XX useful for immobilizing polypeptides via the carboxyl group of the C-
XX terminus. The present sequence represents an immobilised red fluorescent
XX protein variant (Dared) sequence.
SQ Sequence 239 AA;

Query Match          97.7%; Score 1186; DB 8; Length 239;
Best Local Similarity 97.3%; Pred. No. 1.6e-124;
Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRSSKNVIEKEMRKFKVMEGTVNGHEFEIEGEGRPYEGHNTVYKLTXTKGSPLEPFAMDI 60
   |||:|||||
DB 1 MASSENVITEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLTXTKGSPLEPFAMDI 60

QY 61 LSPQFQYSGSKYVYVGPADIPDYKKLSFPEGFKMERVMNFEDEGVVTVTQDSSLQDGCFTY 120
   |||:|||||
DB 61 LSPQFQYSGSKYVYVGPADIPDYKKLSFPEGFKMERVMNFEDEGVVTVTQDSSLQDGCFTY 120

QY 121 KVFIFGVNFPSPDGPVWQKTMGMEASTERYLPKRGVLTGKEIHKALKLKDGGHYLVFESKI 180
   |||:|||||
DB 121 KVFIFGVNFPSPDGPVWQKTMGMEASTERYLPKRGVLTGKEIHKALKLKDGGHYLVFESKI 180

QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   |||:|||||
DB 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 81
ADT77528
ID ADT77528 standard; protein; 239 AA.
XX
AC ADT77528;
XX
DT 13-JAN-2005 (first entry)
XX
DE Protein/linker peptide SEQ ID NO:4 used in protein array.
XX
KW Protein array; planar substrate; immobilisation; detection;
XX target substance; linker peptide.
XX
OS Synthetic.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
XX FT 1..225
XX FT /note= "Corresponds to SEQ ID NO:2"
XX FT 226..239
XX FT /note= "Corresponds to the linker peptide shown on page
XX FT 12 of the specification"
XX
XX WO2004090542-A1.
XX
XX 21-OCT-2004.
XX
XX 09-APR-2004; 2004WO-JP005150.
XX
XX 10-APR-2003; 2003JP-00106450.
XX
XX (NABD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
```

PI Iwakura M, Hirota K;  
 XX  
 DR WPI; 2004-766496/75.  
 XX  
 PT Substrate for protein array, obtained by combining a polymer compound  
 PT having primary amino groups in repeating units on surface of planar  
 PT substrate.  
 PS  
 XX Example 1; SEQ ID NO 4; 32pp; Japanese.  
 CC The invention relates to a substrate for a protein array, obtained by  
 CC combining a polymer compound having primary amino groups in repeating  
 CC units on a surface of a planar substrate. The invention also relates to a  
 CC protein array formed by bonding the C-termini of proteins to the primary  
 CC amino groups of the polymer compound on the substrate. The invention  
 CC further relates to a method for the preparation of the protein array.  
 CC Preferably, the planar substrate has water absorptivity, the polymer used  
 CC is a polyallyl amine or a polylysine, and the immobilised proteins  
 CC contain a linker peptide. The protein array of the invention enables the  
 CC effective and highly sensitive detection of various target substances.  
 CC Sequences ADT7527-ADT7528 represent proteins comprising a C-terminal  
 CC negatively charged acidic linker peptide (ADT7529) which were  
 CC immobilised on a planar substrate via the linker peptide to form a  
 CC protein array.  
 CC  
 CC Sequence 239 AA;  
 SQ  
 Query Match 97.7%; Score 1186; DB 8; Length 239;  
 Best Local Similarity 97.3%; Pred. No. 1,6e-124;  
 Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MSSKNVIEFMRFKRMGTGNGHFEIEGEGEGRPYGHNTVTKLKTGGGLPFPAMD 60  
 DB 1 MASSEAVITEFMRFKRMGTGNGHFEIEGEGEGRPYGHNTVTKLKTGGGLPFPAMD 60  
 QY 61 LSPQFQYGSKVVYKHPADIPDYKLSFPPGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
 DB 61 LSPQFQYGSKVVYKHPADIPDYKLSFPPGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
 QY 121 KYKFTGVNPPSDGPVWQKKTGMGEASTERLYPRDGVLKGEIHAKLKDGGHYLVEFKSI 180  
 DB 121 KYKFTGVNPPSDGPVWQKKTGMGEASTERLYPRDGVLKGEIHAKLKDGGHYLVEFKSI 180  
 QY 181 YMAKKEVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 DB 181 YMAKKEVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 82  
 ABP56686  
 ID ABP56686 standard; protein; 240 AA.  
 XX  
 AC ABP56686;  
 XX  
 DT 25-MAR-2003 (first entry)  
 XX  
 PN 25-MAR-2003 (first entry)  
 XX  
 DE Discosoma red fluorescent protein variant R18H,N24S,F125L and M193K.  
 XX  
 KM Mammalian codon optimised Discosoma red fluorescent protein; variant;  
 KM Discosoma; red fluorescent protein; directed evolution; mutant.  
 XX  
 OS Discosoma sp.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200294992-A2.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 20-MAY-2002; 2002WO-US015968.  
 XX  
 KM 18-MAY-2001; 2001US-0291871P.  
 XX  
 PA (RIGB-) RIGEL PHARM INC.

XX  
 PI Peelle B;  
 XX  
 DR WPI; 2003-120798/11.  
 XX  
 PT New Discosoma red fluorescent protein, useful for functional screens as a  
 PT reporter for gene transcriptions, for target characterization and  
 PT localization of fusion proteins, or for scaffolds for protein and peptide  
 PT libraries.  
 PS  
 XX Example 1; Page; 22pp; English.  
 CC The present invention describes an isolated Discosoma red fluorescent  
 CC protein (1) comprising a 241 residue amino acid sequence (see ABP56678,  
 CC S11), with one or more point mutations at amino acid position N24, F125,  
 CC K164, or M193. Also described: (1) a fusion protein comprising (1); (2)  
 CC an isolated nucleic acid encoding (1); (3) a vector comprising the  
 CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
 CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
 CC encoding (1); and (6) methods of making a fluorescent variant. The  
 CC Discosoma red fluorescent proteins are useful for functional screens as a  
 CC reporter for gene transcriptions (e.g. as a fusion protein), for target  
 CC characterisation and localisation of fusion proteins, or for scaffolds  
 CC for protein and peptide libraries. The fluorescent proteins can also be  
 CC used as selectable markers or reporter molecules for a variety of  
 CC bioassays, including methods that use fluorescence activated cell sorting  
 CC (FACS) as a selection mechanism. The method of directed protein evolution  
 CC is useful for obtaining improved variants of red fluorescent protein. The  
 CC variants of Discosoma red fluorescent proteins have greatly improved  
 CC brightness, expression, and/or folding kinetics as compared to wild type  
 CC or a codon optimised variant. The present sequence represents a mammalian  
 CC codon optimised Discosoma red fluorescent protein variant from the  
 CC present invention. N.B. The present sequence is not given in the  
 CC specification, but is derived from the protein in Fig 1 as stated in Fig  
 CC 2  
 SQ  
 XX Sequence 240 AA;  
 QY  
 Query Match 97.4%; Score 1183; DB 6; Length 240;  
 Best Local Similarity 97.3%; Pred. No. 3,5e-124;  
 Matches 219; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MSSKNVIEFMRFKRMGTGNGHFEIEGEGEGRPYGHNTVTKLKTGGGLPFPAMD 60  
 DB 2 VRSSKNVIEFMRFKRMGTGNGHFEIEGEGEGRPYGHNTVTKLKTGGGLPFPAMD 61  
 QY 61 LSPQFQYGSKVVYKHPADIPDYKLSFPPGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
 DB 62 LSPQFQYGSKVVYKHPADIPDYKLSFPPGFKMERVMNFDGCVTVTODSSLQDGCFTY 121  
 QY 121 KYKFTGVNPPSDGPVWQKKTGMGEASTERLYPRDGVLKGEIHAKLKDGGHYLVEFKSI 180  
 DB 122 KYKFTGVNPPSDGPVWQKKTGMGEASTERLYPRDGVLKGEIHAKLKDGGHYLVEFKSI 181  
 QY 181 YMAKKEVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 DB 182 YMAKKEVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226

RESULT 83  
 ADL26532  
 ID ADL26532 standard; protein; 226 AA.  
 XX  
 AC ADL26532;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 PN 20-MAY-2004 (first entry)  
 XX  
 DE Discosoma fluorescent protein deFP583.  
 XX  
 KM fluorescent protein; fluorescent marker; gene expression;  
 KM protein localization; FP.  
 XX  
 OS Discosoma sp.

```
XX XX MO2004016648-A1.
PN XX
PD XX 26-FEB-2004.
XX XX
XX PF 19-JUL-2003; 2003WO-DE002436.
XX PR 19-JUL-2002; 2002DE-01033082.
XX PA (AMAX-) AMAXA GMBH.
XX PI (AMAX-) AMAXA GMBH.
XX PI Altroge L, Males T;
XX DR MPI; 2004-238793/22.
XX PT New fluorescent proteins, useful as pharmaceuticals, biological markers
PT and for optical information storage, comprise short-lived but restorable
PT fluorescence.
XX PS Example 1; Fig 1; 55pp; German.
XX CC The present invention provides protein and coding sequences of novel
CC fluorescent proteins. These are useful in pharmaceutical compositions and
CC as fluorescent markers for gene expression and protein localization in
CC cellular, developmental or molecular biology, particularly for measuring
CC time-dependent cellular processes such as protein diffusion and transport,
CC also for optical information storage. The present sequence is a Discosoma
CC fluorescent protein.
XX SQ Sequence 226 AA;

Query Match          96.8%; Score 1175.5; DB 8; Length 226;
Best Local Similarity 97.8%; Pred. No. 2.2e-123;
Matches 221; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 MRSSKNVIEKFRKFKVMEGTVNGHEFEIEGEGRPRYEGHNTVTKLKYTKGGPLPFAMD 60
DB 1 MRSSKNVIEKFRKFKVMEGTVNGHEFEIEGEGRPRYEGHNTVTKLKYTKGGPLPFAMD 60
QY 61 LSPFOYGSKYVYVHPADIPDYKKLSFPEGFKMERVWNFEDGGVVTYTOPSSLDGCFIY 120
DB 61 LSPFOYGSKYVYVHPADIPDYKKLSFPEGFKMERVWNFEDGGVVTYTOPSSLDGCFIY 120
QY 121 KVKFTGVNPPSDGPNVQKKTMGWEASTERYPRDGLKGEIHKALKDGGHYLVFERKSI 180
DB 121 KVKFTGVNPPSDGPNVQKKTMGWEASTERYPRDGLKGEIHKALKDGGHYLVFERKSI 180
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225
DB 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 226

RESULT 84
ADC24129
ID ADC24129 standard; protein; 225 AA.
XX
AC ADC24129;
XX
DT 18-DEC-2003 (first entry)
XX
DE Discosoma red fluorescent protein variant fast T1.
XX
XX Discosoma red fluorescent protein; DsRed; AB interface; AC interface;
KW fluorescent protein variant; transcription induction detection;
KW fluorescence energy resonance transfer; FRET; protein kinase;
XX protein phosphatase; ion indicator; mutant; mutein; fast T1.
XX
OS Synthetic.
OS Discosoma.
XX
XX US2003059835-A1.
XX
PD 27-MAR-2003.
```

```
XX XX 10-APR-2002; 2002US-00121258.
PF XX
XX XX 26-FEB-2001; 2001US-00794308.
XX PR 24-MAY-2001; 2001US-0086538.
XX XX
XX PA (TSIE/) TSIE R Y.
XX PA (CAMP/) CAMPBELL R E.
XX PI Tsien RY, Campbell RE;
XX DR MPI; 2003-743764/70.
XX DR N-PSDB; ADC24126.
XX PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein
PT variant having a reduced propensity to oligomerize, useful for detecting
PT transcriptional activity.
XX PS Example 1; SEQ ID NO 4; 67pp; English.
XX CC The invention describes a polynucleotide sequence (I) encoding a
CC Discosoma red fluorescent protein (DsRed) variant having a reduced
CC propensity to oligomerize, comprising amino acid substitutions at the AB
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225
CC amino acids, given in the specification, where the substitutions result
CC in reduced propensity of the DsRed variant to form tetramers. (I) is
CC useful for detecting transcriptional activity by providing a host cells
CC containing a vector which comprises (I) operatively linked to an
CC expression control sequence, and an unit to assay the variant fluorescent
CC protein fluorescence, and assaying fluorescence of the variant
CC fluorescent protein produced by (VII), where variant fluorescent protein
CC fluorescence is indicative of transcriptional activity. A polynucleotide
CC encoding a fusion protein is useful for the analysis of in vivo
CC localization or trafficking of a polypeptide of interest. A polypeptide
CC marker is useful as markers to identify the location and amount of a
CC target protein produced, where the target protein is fused to the marker,
CC as a complement to or alternative for the green fluorescent protein or
CC its spectral variant, for detecting induction of transcription. In
CC applications involving fluorescence energy resonance transfer (FRET),
CC which detects events as the function of the movement of fluorescent
CC donors and acceptors towards or away from each other, for making
CC fluorescent sensors for protein kinase and phosphatase activities or
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the
CC presence of a molecule in a sample, for identifying a specific
CC interaction of a first and second molecule, for determining whether a
CC sample contains an enzyme or for determining the pH of the sample. (I) is
CC useful for identifying a region or condition that regulates the activity
CC of an expression control sequence. This is the amino acid sequence of
CC Discosoma red fluorescent protein variant fast T1.
XX SQ Sequence 225 AA;

Query Match          95.6%; Score 1160; DB 7; Length 225;
Best Local Similarity 96.0%; Pred. No. 1.2e-121;
Matches 216; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRSSKNVIEKFRKFKVMEGTVNGHEFEIEGEGRPRYEGHNTVTKLKYTKGGPLPFAMD 60
DB 1 MRSSKNVIEKFRKFKVMEGTVNGHEFEIEGEGRPRYEGHNTVTKLKYTKGGPLPFAMD 60
QY 61 LSPFOYGSKYVYVHPADIPDYKKLSFPEGFKMERVWNFEDGGVVTYTOPSSLDGCFIY 120
DB 61 LSPFOYGSKYVYVHPADIPDYKKLSFPEGFKMERVWNFEDGGVVTYTOPSSLDGCFIY 120
QY 121 KVKFTGVNPPSDGPNVQKKTMGWEASTERYPRDGLKGEIHKALKDGGHYLVFERKSI 180
DB 121 KVKFTGVNPPSDGPNVQKKTMGWEASTERYPRDGLKGEIHKALKDGGHYLVFERKSI 180
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225
DB 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225
```



```
RESULT 85
ADL46206
ID ADL46206 standard; protein; 225 AA.
XX
AC ADL46206;
XX
DT 20-MAY-2004 (first entry)
XX
DE Discosoma red fluorescent protein (Dared) variant T1 protein.
XX
KM red fluorescent protein; Dared; fluorescence; red wavelength;
XX oligomerization; tetramerization; immunoassay; hybridization assay.
XX
OS Discosoma sp.
XX
PN MO2003086446-A1.
XX
PD 23-OCT-2003.
XX
PF 09-APR-2003; 2003WO-US010879.
XX
PR 10-APR-2002; 2002US-00121258.
XX 29-JUL-2002; 2002US-00209208.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Teien RY, Campbell RE, Baird GS;
XX
DR WPI; 2003-845265/78.
XX N-PSDB; ADL46207.
XX
PT New monomeric and dimeric Anthozoan fluorescent protein variants with
PT reduced propensity to oligomerize, and encoding polynucleotides, useful
PT in molecular biology, e.g. in immunoassays or in tracking protein
PT movement in cells.
XX
PS Disclosure; SEQ ID NO 4; 166pp; English.
XX
XX
CC The invention relates to a polynucleotide sequence encoding a Discosoma
CC red fluorescent protein (Dared) variant having a reduced propensity to
CC oligomerize. The protein variant comprises one or more amino acid
CC substitutions at the AB and/or AC interface(s) of the wild-type Dared
CC sequence, where the substitutions result in reduced propensity of the
CC Dared variant to form tetramers and where the variant displays detectable
CC fluorescence of at least one red wavelength. The composition and methods
CC are useful in producing red fluorescent proteins having reduced
CC propensity for oligomerization, especially tetramerization. The protein
CC may be used in molecular biology and in other scientific applications,
CC such as in immunoassays or hybridization assays, or in tracking the
CC movement of proteins in cells. This sequence corresponds to the Dared
CC variant T1 protein.
XX
SQ Sequence 225 AA:

Query Match          95.6%; Score 1160; DB 7; Length 225;
Best Local Similarity 96.0%; Pred. No. 1.2e-121;
Matches 216; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRSSKNVYKEFPMFKRMETGVNGHFEIEGEGGRPYEGHNTVKLVTKGGLPFPAMD1 60
DB 1 MASSEDIKEFMRKVMGSGVNGHFEIEGEGGRPYEGTQAKLVTKGGLPFPAMD1 60
QY 61 LSPQFOYGSKVYVYKHADIPDYKLSFPEGFKMERVNNFEDGAVVTYTDSSIQDGCFTY 120
DB 61 LSPQFOYGSKVYVYKHADIPDYKLSFPEGFKMERVNNFEDGAVVTYTDSSIQDGSFTY 120
QY 121 KVKFIGNVPSDGPVWQKKTWGEASTERLYPRDGVLKGEIHKALKDKGGHYLVERFSI 180
DB 121 KVKFIGNVPSDGPVWQKKTWGEASTERLYPRDGVLKGEIHKALKDKGGHYLVERFSI 180
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRRHFL 225
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRRHFL 225
```

```
RESULT 86
ADL46226
ID ADL46226 standard; protein; 225 AA.
XX
AC ADL46226;
XX
DT 20-MAY-2004 (first entry)
XX
DE Discosoma red fluorescent protein (Dared) T1 protein with I125R mutation.
XX
KM red fluorescent protein; Dared; fluorescence; red wavelength;
XX oligomerization; tetramerization; immunoassay; hybridization assay.
XX
OS Discosoma sp.
XX
PN MO2003086446-A1.
XX
PD 23-OCT-2003.
XX
PF 09-APR-2003; 2003WO-US010879.
XX
PR 10-APR-2002; 2002US-00121258.
XX 29-JUL-2002; 2002US-00209208.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Teien RY, Campbell RE, Baird GS;
XX
DR WPI; 2003-845265/78.
XX
PT New monomeric and dimeric Anthozoan fluorescent protein variants with
PT reduced propensity to oligomerize, and encoding polynucleotides, useful
PT in molecular biology, e.g. in immunoassays or in tracking protein
PT movement in cells.
XX
PS Disclosure; SEQ ID NO 24; 166pp; English.
XX
XX
CC The invention relates to a polynucleotide sequence encoding a Discosoma
CC red fluorescent protein (Dared) variant having a reduced propensity to
CC oligomerize. The protein variant comprises one or more amino acid
CC substitutions at the AB and/or AC interface(s) of the wild-type Dared
CC sequence, where the substitutions result in reduced propensity of the
CC Dared variant to form tetramers and where the variant displays detectable
CC fluorescence of at least one red wavelength. The composition and methods
CC are useful in producing red fluorescent proteins having reduced
CC propensity for oligomerization, especially tetramerization. The protein
CC may be used in molecular biology and in other scientific applications,
CC such as in immunoassays or hybridization assays, or in tracking the
CC movement of proteins in cells. This sequence corresponds to the Dared
CC variant T1 protein with a I125R mutation.
XX
SQ Sequence 225 AA:

Query Match          95.0%; Score 1153; DB 7; Length 225;
Best Local Similarity 95.6%; Pred. No. 7.6e-121;
Matches 215; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRSSKNVYKEFPMFKRMETGVNGHFEIEGEGGRPYEGHNTVKLVTKGGLPFPAMD1 60
DB 1 MASSEDIKEFMRKVMGSGVNGHFEIEGEGGRPYEGTQAKLVTKGGLPFPAMD1 60
QY 61 LSPQFOYGSKVYVYKHADIPDYKLSFPEGFKMERVNNFEDGAVVTYTDSSIQDGCFTY 120
DB 61 LSPQFOYGSKVYVYKHADIPDYKLSFPEGFKMERVNNFEDGAVVTYTDSSIQDGSFTY 120
QY 121 KVKFIGNVPSDGPVWQKKTWGEASTERLYPRDGVLKGEIHKALKDKGGHYLVERFSI 180
DB 121 KVKFIGNVPSDGPVWQKKTWGEASTERLYPRDGVLKGEIHKALKDKGGHYLVERFSI 180
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRRHFL 225
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRRHFL 225
```

Db 181 YMAKKPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERABGRHHLFL 225

RESULT 87

ADM13537

ADM13537 standard; protein; 225 AA.

XX

AC ADM13537;

XX

DT 15-JUL-2004 (first entry)

XX

DE Superfolding coral Dered protein #1.

XX

KM GFP; superfolder mutant; mutein; green fluorescent protein;

XX

KM directed-evolution; folding interference domain; reporter protein;

XX

mutant; Dered; coral.

XX

OS Diacosoma sp.

XX

OS Synthetic.

XX

PN US2004078146-A1.

XX

XX 22-APR-2004.

XX

PF 24-APR-2003; 2003US-00423688.

XX

PR 24-APR-2002; 2002US-00132067.

XX

PA (WALD/) WALDO G S.

XX

PI Waldo GS;

XX

DR MPI: 2004-340059/31.

DR N-FSDB; ADM13536.

XX

PT New modified green fluorescent protein, useful as a reporter in

XX

expression studies.

XX

PS Claim 18; SEQ ID NO 41; 46pp; English.

XX

CC The invention relates to a green fluorescent protein (GFP) that comprises

CC at least 80% identity to the GFP superfolding mutant appearing as

CC ADM13535 that has at least one amino acid substitution selected from

CC e.g., a substitution at position 30 that is an arginine or a conservative

CC variant of arginine and measurable fluorescence activity. Also included

CC are directed-evolution for generating an enhanced folding variant of a

CC target polypeptide (comprising mutating a polynucleotide encoding

CC polypeptide of interest to generate a library of mutated polynucleotides,

CC linking mutated polynucleotides to polynucleotide encoding folding

CC interference domain to form fusion protein (FP) constructs, expressing

CC FP, and selecting FP that display optimal folding activity in relation to

CC FP comprising wild-type polypeptide and folding interference domain, thus

CC identifying polypeptide with enhanced folding activity), an enhanced

CC folding variant of a fluorescent protein generated by the method, an

CC enhanced folding variant of a chromophoric protein generated by the

CC method, enhancing folding of a polypeptide comprising multiple domains

CC (involving joining a first domain of the polypeptide to a poorly folding

CC domain, to form a fusion protein, mutating the first domain, detecting an

CC increase in the amount of activity generated by a first mutated fusion

CC protein in comparison to a fusion protein comprising a wild-type first

CC domain and the poorly folding polypeptide domain, thus identifying a

CC first domain with enhanced folding, joining a second domain of the

CC polypeptide to the first mutated fusion protein to form a second fusion

CC protein, mutating the second domain, and detecting an increase in the

CC amount of activity generated by a second mutated fusion protein in

CC comparison to a fusion protein comprising the wild-type second domain and

CC the first mutated fusion protein, thus identifying a target polypeptide

CC with multiple domains that have enhanced folding), an isolated nucleic

CC acid encoding the GFP, an expression vector comprising the nucleic acid

CC and a host cell comprising the vector. The Dered protein from coral is

CC also considered for mutation using the above method. The enhanced folding

CC variants are useful as reporter proteins to express the report level of a

CC protein. The method is useful for directed-evolution of generating

CC enhanced folding variant of target polypeptide and for improving folding

CC and solubility of a target protein. The present sequence represents a

CC coral Dered enhanced folding variant. NOTE: The authors refer to SEQ ID

CC 38-42 as SEQ ID 1-5 throughout the specification and never refer to SEQ

CC ID 1-37 by number, it is obvious that SEQ ID 38-42 are meant to be SEQ ID

CC 1-5.

XX

XX Sequence 225 AA:

XX

Query Match 93.2%; Score 1132; DB 8; Length 225;

Best Local Similarity 92.4%; Pred. No. 1,86-118;

Matches 208; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 MRSSKNVTKEMRRKRVNMEGTVMNHEPEIIEGEGRPPEHNTYKLTXTKGAPLPFADI 60

Db 1 MESSSDVTKEMRRKRVNMEGTVNHEPEIIEGEGRPPEHNTYKLTXTKGAPLPFADI 60

QY 61 LSPQFGSKYVYVHGPADIPYKLLSPFEGFKWERVNNFEDGVVTVTQDSSLQDGCPIY 120

Db 61 LSPQFGSKYVYVHGPADIPYKLLSPFEGFKWERVNNFEDGVVTVTQDSSLQDGCPIY 120

QY 121 KYKFIGVNFPSDGPVMQKKTGMWASTERLYPRDGVLYKGEIHKALKLKDGGHYLVFESI 180

Db 121 KYKFIGVNFPSDGPVMQKKTGMWASTERLYPRDGVLYKGEIHKALKLKDGGHYLVDIRSI 180

QY 181 YMAKKPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

Db 181 YMAKKPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERABGRHHLFL 225

RESULT 88

ADC24131

ID ADC24131 standard; protein; 226 AA.

XX

AC ADC24131;

XX

DT 18-DEC-2003 (first entry)

XX

DE Diacosoma red fluorescent protein variant dimer2.

XX

XX

XX Diacosoma red fluorescent protein; Dered; AB interface; AC interface;

XX

KM fluorescent protein variant; transcription induction detection;

XX

KM fluorescence energy resonance transfer; FRPT; protein kinase;

XX

XX protein phosphatase; ion indicator; mutant; dimer; mutein.

OS

OS Synthetic.

OS Diacosoma.

XX

PN US2003059835-A1.

XX

PD 27-MAR-2003.

XX

PF 10-APR-2002; 2002US-00121258.

XX

PR 26-FEB-2001; 2001US-00794308.

XX

PR 24-MAY-2001; 2001US-00866538.

XX

PA (TSIE/) TSIEH R Y.

XX

PA (CAMP/) CAMPBELL R E.

XX

PI Tsien RY, Campbell RE;

XX

DR MPI: 2003-743764/70.

DR N-FSDB; ADC24130.

XX

XX

XX Novel polynucleotide sequence encoding Diacosoma red fluorescent protein

XX

XX variant having a reduced propensity to oligomerize, useful for detecting

XX

XX transcriptional activity.

PS

PS Claim 12; SEQ ID NO 6; 67pp; English.

XX

CC The invention describes a polynucleotide sequence (I) encoding a

CC Diacosoma red fluorescent protein (DsRed) variant having a reduced

propensity to oligomerize, comprising amino acid substitutions at the AB and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225 amino acids, given in the specification, where the substitutions result in reduced propensity of the DsRed variant to form tetramers. (1) is useful for detecting transcriptional activity by providing a host cells containing a vector which comprises (i) operatively linked to an expression control sequence, and an unit to assay the variant fluorescent protein fluorescence, and assaying fluorescence of the variant fluorescent protein produced by (VII), where variant fluorescent protein fluorescence is indicative of transcriptional activity. A polynucleotide encoding a fusion protein is useful for the analysis of in vivo localization or trafficking of a polypeptide of interest. A polypeptide marker is useful as markers to identify the location and amount of a target protein produced, where the target protein is fused to the marker, as a complement to or alternative for the green fluorescent protein or its spectral variant, for detecting induction of transcription, in applications involving fluorescence energy resonance transfer (FRET), which detects events as the function of the movement of fluorescent donors and acceptors towards or away from each other, for making fluorescent sensors for protein kinase and phosphatase activities or indicators for ions and molecules such as  $Ca^{2+}$ ,  $Zn^{2+}$ , for identifying the presence of a molecule in a sample, for identifying a specific interaction of a first and second molecule, for determining whether a sample contains an enzyme or for determining the pH of the sample. (1) is useful for identifying a region or condition that regulates the activity of an expression control sequence. This is the amino acid sequence of Discosoma red fluorescent protein variant dimer2.

Sequence 226 AA;

Query Match 92.3%; Score 1121; DB 7; Length 226;

Best Local Similarity 92.8%; Pred. No. 3.1e-117;

Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

```

QY 3 SSKNTIKEMRKVRKMEGVNHEFEIIEGSGRPRYEGNNTYKLVTKGGRPLPFAMDILS 62
   ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 4 SSEDVTKEMRFRVMEGVNHEFEIIEGSGRPRYEGNNTYKLVTKGGRPLPFAMDILS 63
QY 63 POFQVGSKAYVGHADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSLDGCFIYKY 122
   ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 64 POFQVGSKAYVGHADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSLDGTLIYKY 123
QY 123 KFIGVNFPSDGPVMOQKKTGWGEASTERLYPRDGLKGEIHAKLKDGGHYLVEFKSIYM 182
   ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 124 KFRGNTFPPDGPVMOQKKTGWGEASTERLYPRDGLKGEIHAKLKDGGHYLVEFKSIYM 183
QY 183 AKKPVQLPGYYVVDTKLDTSHNEDYTYVEQYERTEGRHHLFL 225
   ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 184 AKKPVQLPGYYVVDTKLDTSHNEDYTYVEQYERTEGRHHLFL 226

```

RESULT 89

ADL46208 ID ADL46208 standard; protein; 226 AA.

AC ADL46208;

DT 20-MAY-2004 (first entry)

DE Discosoma red fluorescent protein (DsRed) variant dimer2 protein.

KM red fluorescent protein; DsRed; fluorescence; red wavelength; oligomerization; tetramerization; immunoassay; hybridization assay.

OS Discosoma sp.

PN WO2003086446-A1.

PD 23-OCT-2003.

PF 09-APR-2003; 2003MO-US010879.

PR 10-APR-2002; 2002US-00121258.

PR 29-JUL-2002; 2002US-00209208.

XX (REGC ) UNIV CALIFORNIA.

PI Tsien RY, Campbell RE, Baird GS;

DR WPI, 2003-845265/78.

XX N-PSDB; ADL46209.

PT New monomeric and dimeric Anthozoan fluorescent protein variants with reduced propensity to oligomerize, and encoding polynucleotides, useful in molecular biology, e.g. in immunoassays or in tracking protein movement in cells.

PS Disclosure; SEQ ID NO 6; 166pp; English.

The invention relates to a polynucleotide sequence encoding a Discosoma red fluorescent protein (DsRed) variant having a reduced propensity to oligomerize. The protein variant comprises one or more amino acid substitutions at the AB and/or AC interface(s) of the wild-type DsRed sequence, where the substitutions result in reduced propensity of the DsRed variant to form tetramers and where the variant displays detectable fluorescence of at least one red wavelength. The composition and methods are useful in producing red fluorescent proteins having reduced propensity for oligomerization, especially tetramerization. The protein may be used in molecular biology and in other scientific applications, such as in immunoassays or hybridization assays, or in tracking the movement of proteins in cells. This sequence corresponds to the DsRed variant dimer2 protein.

Sequence 226 AA;

Query Match 92.3%; Score 1121; DB 7; Length 226;

Best Local Similarity 92.8%; Pred. No. 3.1e-117;

Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

```

QY 3 SSKNTIKEMRKVRKMEGVNHEFEIIEGSGRPRYEGNNTYKLVTKGGRPLPFAMDILS 62
   ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 4 SSEDVTKEMRFRVMEGVNHEFEIIEGSGRPRYEGNNTYKLVTKGGRPLPFAMDILS 63
QY 63 POFQVGSKAYVGHADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSLDGCFIYKY 122
   ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 64 POFQVGSKAYVGHADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSLDGTLIYKY 123
QY 123 KFIGVNFPSDGPVMOQKKTGWGEASTERLYPRDGLKGEIHAKLKDGGHYLVEFKSIYM 182
   ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 124 KFRGNTFPPDGPVMOQKKTGWGEASTERLYPRDGLKGEIHAKLKDGGHYLVEFKSIYM 183
QY 183 AKKPVQLPGYYVVDTKLDTSHNEDYTYVEQYERTEGRHHLFL 225
   ||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 184 AKKPVQLPGYYVVDTKLDTSHNEDYTYVEQYERTEGRHHLFL 226

```

RESULT 90

AAI99838 ID AAY99838 standard; protein; 225 AA.

AC AAY99838;

DT 19-SEP-2000 (first entry)

DE Humanised anthozoa novel fluorescent protein hybrid drFP563/dmFP592.

XX Anthozoa; drFP563; fluorescent protein; non-bioluminescent organism; fluorescent labeling; hybrid.

OS Discosoma sp; "red".

OS Synthetic.

PN WO200034326-A1.

PD 15-JUN-2000.

```

PF 10-DEC-1999; 99WO-US029473.
XX
PR 11-DEC-1998; 98US-00210330.
PR 14-OCT-1999; 99US-00418529.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;
PI Ding L;
XX WPI; 2000-423381/36.
XX DR N-PSDB; AAA48748.
XX
XX Novel fluorescent protein from non-bioluminescent Discosoma sp. red,
XX useful for fluorescent labeling and as markers.
XX
XX Claim 20; Page 78; 86pp; English.
XX
XX The present sequence is humanised drfp583/dmfp592. Amplified wild-type
XX coding region fragments from drfp583 and dmfp592 were digested with
XX restriction endonucleases and combined together to produce a hybrid
XX construct. drfp583 is a full-length cDNA encoding a novel fluorescent
XX protein (mFP) from Discosoma sp. "red", a non-bioluminescent species of
XX the Class Anthozoa. The hybrid sequence was changed to optimise the
XX codons for expression of the protein in mammalian cells. Fluorescent
XX proteins can be used in fluorescent labeling, a useful tool for marking a
XX protein, cell or organism of interest. Unlike other markers used in
XX protein labeling, such as beta-galactosidase and luciferase, fluorescent
XX proteins do not require an exogenous cofactor or substrate. Methods
XX involving fluorescent proteins are also less laborious and less difficult
XX to control than the traditional methods of fluorescent labeling, where a
XX protein of interest is purified and then covalently conjugated to a
XX fluorophore derivative. Novel fluorescent proteins isolated from species
XX of the Class Anthozoa can be used as markers for gene expression and
XX protein localization studies, and in fluorescence resonance energy
XX transfer (FRET) reactions. They may have improved properties and better
XX suitability for larger excitations compared to prior art fluorescent
XX proteins such as green fluorescent protein
XX
XX Sequence 225 AA;
SQ

```

Query Match 92.2%; Score 1119; DB 3; Length 225;  
 Best Local Similarity 91.6%; Pred. No. 5.1e-117;  
 Matches 206; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 MRSKNVYKEFMRRKVMRGTNGHGFPIBEGGGRPYEGHNTYKLTKTGGPLPFPADI 60
   |||||
DB 1 MSCSNVYKEFMRRKVMRGTNGHGFPIBEGGGRPYEGHNTYKLTKTGGPLPFPADI 60
QY 1 LSPFOYGSKYVYVHPADIPIYKLTSPPEGFKMERVMNFPEDGVTYVYQDSSLQDGCFTY 120
   |||||
DB 61 LSPFOYGSKYVYVHPADIPIYKLTSPPEGFKMERVMNFPEDGVTYVYQDSSLQDGCFTY 120
QY 121 KVFIFGNVFPSPDGVPMQKKTGMWEASSTERLYPRDGLKGEIHKALDKDGGHYLVEFKSI 180
   |||||
DB 121 EVKFIGVNFPSDGVPMQKKTGMWEASSTERLYPRDGLKGDIMALRLRGGGHYLVVEFKSI 180
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTYVQYERTEGRHHLFL 225
   |||||
DB 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTYVQYERTEGRHHLFL 225

```

RESULT 91  
 ID AAE28835 standard; protein; 225 AA.  
 XX AAE28835;  
 XX  
 DT 27-DEC-2002 (first entry)  
 XX  
 XX *Discosoma* sp. drfp583/dmfp592 (6/90) protein.  
 DE  
 XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;

```

KW fluorescence activated cell sorting application; fluorescent timer;
KW biosensor; fluorescence resonance energy transfer application; FRET;
KW colouring agent; recombinant DNA application; analyte detection assay;
KW sunscreen; second messenger detector; drfp583 protein; dmfp592 protein.
XX
XX Discosoma sp.
XX
XX WO200268459-A2.
XX
XX 06-SEP-2002.
XX
XX 20-FEB-2002; 2002WO-US005749.
XX
XX 21-FEB-2001; 2001US-0270983P.
XX PR 04-DEC-2001; 2001US-00006922.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;
XX
XX WPI; 2002-691654/74.
XX DR N-PSDB; AAD46280.
XX
XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant
XX of an aggregating Cnidarian chromo- or fluorescent protein or mutant for
XX analyte detection assays or fluorescence activated cell sorting
XX applications.
XX
XX Disclosure; Page 72-73; 80pp; English.
XX
XX The invention relates to nucleic acid molecules encoding non-aggregating
XX chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are
XX useful in analyte detection assays, as colouring agents, as markers in
XX recombinant DNA applications, as sunscreens or filters, in fluorescence
XX resonance energy transfer (FRET) applications, as biosensors in
XX prokaryotic and eukaryotic cells, in screening assays, as second
XX messenger detectors, in fluorescence activated cell sorting applications,
XX in protease cleavage assays or as fluorescent timers. The present
XX sequence is Discosoma sp. drfp583/dmfp592 (6/90) protein of the invention
XX
XX Sequence 225 AA;
SQ

```

Query Match 92.2%; Score 1119; DB 5; Length 225;  
 Best Local Similarity 91.6%; Pred. No. 5.1e-117;  
 Matches 206; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 MRSKNVYKEFMRRKVMRGTNGHGFPIBEGGGRPYEGHNTYKLTKTGGPLPFPADI 60
   |||||
DB 1 MSCSNVYKEFMRRKVMRGTNGHGFPIBEGGGRPYEGHNTYKLTKTGGPLPFPADI 60
QY 1 LSPFOYGSKYVYVHPADIPIYKLTSPPEGFKMERVMNFPEDGVTYVYQDSSLQDGCFTY 120
   |||||
DB 61 LSPFOYGSKYVYVHPADIPIYKLTSPPEGFKMERVMNFPEDGVTYVYQDSSLQDGCFTY 120
QY 121 KVFIFGNVFPSPDGVPMQKKTGMWEASSTERLYPRDGLKGEIHKALDKDGGHYLVEFKSI 180
   |||||
DB 121 EVKFIGVNFPSDGVPMQKKTGMWEASSTERLYPRDGLKGDIMALRLRGGGHYLVVEFKSI 180
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTYVQYERTEGRHHLFL 225
   |||||
DB 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTYVQYERTEGRHHLFL 225

```

RESULT 92  
 ID AAY99839 standard; protein; 225 AA.  
 XX AAY99839;  
 XX  
 DT 19-SEP-2000 (first entry)  
 XX  
 XX Humanised Anthozoa novel fluorescent protein hybrid drfp583/dmfp592-2G.  
 DE  
 XX

KW Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
KW fluorescent labeling; hybrid; mutant; mutein.  
XX  
XX  
OS Diacosoma sp; "red".  
OS Synthetic.  
XX  
XX WO20034326-A1.  
XX  
PD 15-JUN-2000.  
XX  
XX 10-DEC-1999; 99WO-US029473.  
XX  
XX 11-DEC-1998; 98US-00210330.  
PR 14-OCT-1999; 99US-00418529.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
XX Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
PI Ding L;  
XX  
XX MPI: 2000-423381/36.  
DR N-ESDB; AAA48749.  
XX  
XX Novel fluorescent protein from non-bioluminescent Diacosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
XX  
XX Claim 6; Page 79; 86pp; English.  
XX  
XX The present sequence is derived from humanised hybrid protein  
CC drFP583/dmFP592. drFP583 is a full-length cDNA encoding a novel  
CC fluorescent protein (nFP) from Diacosoma sp. "red", a non-bioluminescent  
CC species of the Class Anthozoa. The hybrid sequence was changed to  
CC optimize the codons for expression of the protein in mammalian cells. The  
CC present sequence contains two substitutions, K15Q and T217S, which were  
CC incorporated by random mutagenesis. Fluorescent proteins can be used in  
CC fluorescent labeling, a useful tool for marking a protein, cell or  
CC organism of interest. Unlike other markers used in protein labeling, such  
CC as beta-galactosidase and luciferase, fluorescent proteins do not require  
CC an exogenous cofactor or substrate. Methods involving fluorescent  
CC proteins are also less laborious and less difficult to control than the  
CC traditional methods of fluorescent labeling, where a protein of interest  
CC is purified and then covalently conjugated to a fluorophore derivative.  
CC Novel fluorescent proteins isolated from species of the Class Anthozoa  
CC can be used as markers for gene expression and protein localization  
CC studies, and in fluorescence resonance energy transfer (FRET) reactions.  
CC They may have improved properties and better suitability for larger  
CC excitations compared to prior art fluorescent proteins such as green  
CC fluorescent protein  
XX  
XX Sequence 225 AA;  
SQ  
Query Match 91.5%; Score 1111; DB 3; Length 225;  
Best Local Similarity 90.7%; Pred. No. 4.1e-116;  
Matches 204; Conservative 14; Mismatches 7; Indels 0; Gaps 0;  
QY 1 MSSSKVIVKFMKFRKMGSTVNGHFEIIGEGEGRPYEGSHNTVTKLVTGKGLPPAFMDI 60  
DB 1 MSCSKVIVKFMKFRKMGSTVNGHFEIIGEGEGRPYEGSHNTVTKLVTGKGLPPAFMDI 60  
QY 61 LSPQFOYGSKVYVKHAPADIPDYKGLSPPEGFKMERVMNPFEDGCVTVTODSSLKDCGCFIY 120  
DB 61 LSPQFOYGSKVYVKHAPADIPDYKGLSPPEGFKMERVMNPFEDGCVTVTODSSLKDCGCFIY 120  
QY 121 KVFVIGVNFPSDGPVWQKKTWGWASERLYPRDGYLKGSIHKLKXKDGSHLVPRKSI 180  
DB 121 EYVFVIGVNFPSDGPVWQKKTWGWASERLYPRDGYLKGSIHKLKXKDGSHLVPRKSI 180  
QY 181 YMAKKPVQLPGYVYVDSKDIITSHNEDYTIIVEOYERTGRRHLLFL 225  
DB 181 YMAKKPVQLPGYVYVDSKDIITSHNEDYTIIVEOYERTGRRHLLFL 225

AAE28928  
ID AAE28928 standard; protein; 225 AA.  
XX  
XX  
AC AAE28928;  
XX  
XX 27-DEC-2002 (first entry)  
XX  
XX Diacosoma sp. drFP583/dmFP592 (6/90) mutant protein (K15Q+T217S).  
XX  
XX  
XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; dmFP592 protein;  
KW mutant; mutein.  
XX  
XX Diacosoma sp.  
OS Synthetic.  
XX  
XX  
XX Key Location/Qualifiers  
XX  
XX Misc-difference 15  
FT /note= "Wild-type Lys substituted with Gln"  
FT 217  
FT Misc-difference 217  
FT /note= "Wild-type Thr substituted with Ser"  
XX  
XX WO200268459-A2.  
XX  
XX 06-SEP-2002.  
XX  
XX 20-FEB-2002; 2002WO-US005749.  
XX  
XX 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
XX Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
PI MPI: 2002-691654/74.  
XX  
XX  
XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Crdarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
XX  
XX Disclosure; Page; 80pp; English.  
XX  
XX  
XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Diacosoma sp. drFP583/dmFP592 (6/90) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Diacosoma sp. drFP583/dmFP592 (6/90) wild-type protein shown  
CC as SEQ ID NO:12 (AAE28835) in page 72-73 of the specification  
XX  
XX Sequence 225 AA;  
SQ  
Query Match 91.5%; Score 1111; DB 5; Length 225;  
Best Local Similarity 90.7%; Pred. No. 4.1e-116;  
Matches 204; Conservative 14; Mismatches 7; Indels 0; Gaps 0;  
QY 1 MSSSKVIVKFMKFRKMGSTVNGHFEIIGEGEGRPYEGSHNTVTKLVTGKGLPPAFMDI 60  
DB 1 MSCSKVIVKFMKFRKMGSTVNGHFEIIGEGEGRPYEGSHNTVTKLVTGKGLPPAFMDI 60  
QY 61 LSPQFOYGSKVYVKHAPADIPDYKGLSPPEGFKMERVMNPFEDGCVTVTODSSLKDCGCFIY 120  
DB 61 LSPQFOYGSKVYVKHAPADIPDYKGLSPPEGFKMERVMNPFEDGCVTVTODSSLKDCGCFIY 120

QY	121	KYKFFGVNFPSPGPMQKKTGMEASTERYLPDGYLKGEIHAKLKXGGHYLVFPKSI	180
DB	121	EVKFFGVNFPSPGPMQKKTGMEASTERYLPDGYLKGEIHAKLKXGGHYLVFPKSI	180
QY	181	YMAKFPVOLPGYYVDSKLDITSHNEDYITVEQYKTEGRHHLFL	225
DB	181	YMAKFPVOLPGYYVDSKLDITSHNEDYITVEQYKTEGRHHLFL	225
RESULT 94			
ID	AAE28929	standard; protein; 223 AA.	
XX	AAE28929;		
DT	27-DEC-2002	(first entry)	
DE	Discoosoma sp. drfp583/dmfp592 (6/90)	mutant protein (Sdel+Cdel+K5E+K9T).	
KM	Fluorescent protein; chromoprotein; protease cleavage assay; filter;		
KM	fluorescence activated cell sorting application; fluorescent timer;		
KM	biosensor; fluorescence resonance energy transfer application; FRET;		
KM	colouring agent; recombinant DNA application; analyte detection assay;		
KM	sunscreen; second messenger detector; drfp583 protein; dmfp592 protein;		
KM	mutant; mutein.		
OS	Discoosoma sp.		
XX	Synthetic.		
XX			
FT	Key	Location/Qualifiers	
FT	Misc-difference 5	/note= "Wild-type Lys substituted with Glu"	
FT	Misc-difference 9	/note= "Wild-type Lys substituted with Thr"	
XX	WO200268459-A2.		
XX	06-SEP-2002.		
PD	20-FEB-2002; 2002MO-US005749.		
XX	21-FEB-2001; 2001US-0270983P.		
PR	04-DEC-2001; 2001US-00006922.		
XX	(CLON-) CLONTECH LAB INC.		
XX	Lukyanov S, Lukyanov K, Vannukhovich Y, Savitsky A, Fradkov A;		
PI	WPI; 2002-691654/74.		
DR			
XX			
PT	New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant		
PT	of an aggregating Candarian chromo- or fluorescent protein or mutant for		
PT	analyte detection assays or fluorescence activated cell sorting		
PT	applications.		
XX			
XX	Disclosure; Page; 80pp; English.		
XX			
CC	The invention relates to nucleic acid molecules encoding non-aggregating		
CC	chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are		
CC	useful in analyte detection assays, as colouring agents, as markers in		
CC	recombinant DNA applications, as sunscreens or filters, in fluorescence		
CC	resonance energy transfer (FRET) applications, as biosensors in		
CC	prokaryotic and eukaryotic cells, in screening assays, as second		
CC	messenger detectors, in fluorescence activated cell sorting applications,		
CC	in protease cleavage assays or as fluorescent timers. The present		
CC	sequence is Discoosoma sp. drfp583/dmfp592 (6/90) mutant protein of the		
CC	invention. Note: This sequence is not shown in the specification, but is		
CC	derived from Discoosoma sp. drfp583/dmfp592 (6/90) wild-type protein shown		
CC	as SEQ ID NO:12 (AAE28935) in page 72-73 of the specification		
XX	Sequence 223 AA;		

	Query Match	91.1%, Score 1106; DB 5; Length 223;
	Best Local Similarity	91.4%, Pred. No. 1,5e-115;
	Matches 203; Conservative	13; Mismatches 6; Indels 0; Gaps 0;
OY	4 SKNYIKEMRKRYMEGTGVNGHFEIEEGEGRGREYGHNTVLTAKTKTGKGLPFPAMDILSP	63
Dd	2 SENVITTEPMRKVMNEGTVNQHFEIENGSGGRGYBHGCSVKLMTWKGGPLPPAFADILSP	61
OY	64 QFOYGSKYTVKHDPDIDPYKKLSFPEGFKMYRRVNANFEDGCVTYTVDSSLDGCCTITYKVK	123
Dd	62 QFOYGSKYTVKHDPDIDPYKKLSFPEGFKMYRVANNFBDGCVTYVSQDSLSLDGCCTITYEVK	121
OY	124 FIGNPFDEGDGMOKTMGEASTERYLPYRPGVLKGELHKALKDKDGSHLYVEFSITYMA	183
Dd	122 FIGNPFDEGDGMQRIRMGFAASSERLYPRGVLKGDIMLRLEGGSHLYVEFSITYMA	181
OY	184 KKPVOLPGYYVDSKLDTISTHNEDYTIVEOYERTEGRHHLFL	225
Dd	182 KKPVOLPGYYVDSKLDTISTHNEDYTIVEOYERTEGRHHLFL	223

RESULT 95  
AA99840  
ID AA99840 standard; protein; 225 AA.  
XX  
XX  
AC  
XX  
AA99840;  
XX  
DT 19-SEP-2000 (first entry)  
XX  
XX Humanised Anthozoa novel fluorescent protein hybrid drfp583/dmfp592-03.  
XX  
XX Anthozoa; drfp583; fluorescent protein; non-bioluminescent organism;  
KW fluorescent labeling; hybrid; mutant; mutein.  
XX  
XX Discosoma sp; "red".  
OS Synthetic.  
XX  
XX MO200034326-A1.  
XX  
XX  
PD 15-JUN-2000.  
XX  
XX  
PP 10-DEC-1999; 99WO-US029473.  
XX  
XX 11-DEC-1998; 98US-00210330.  
PR 14-OCT-1999; 99US-00418529.  
XX  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
PI Ding L;  
XX  
XX WPI; 2000-423381/36.  
DR  
DR N-PSDB; AAA48750.  
XX  
XX Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
XX  
XX  
PS Claim 20; Page 81-82; 86pp; English.  
XX  
XX  
CC The present sequence is derived from humanised hybrid protein  
CC drfp583/dmfp592. drfp583 is a full-length cDNA encoding a novel  
CC fluorescent protein (nfp) from Discosoma sp. "red", a non-bioluminescent  
CC species of the Class Anthozoa. The hybrid sequence was changed to  
CC optimise the codons for expression of the protein in mammalian cells. The  
CC present sequence contains three substitutions, K15Q, K83M and T217S,  
CC which were incorporated by random mutagenesis. Fluorescent proteins can  
CC be used in fluorescent labeling, a useful tool for marking a protein,  
CC cell or organism of interest. Unlike other markers used in protein  
CC labeling, such as beta-galactosidase and luciferase, fluorescent proteins  
CC do not require an exogenous cofactor or substrate. Methods involving  
CC fluorescent proteins are also less laborious and less difficult to  
CC control than the traditional methods of fluorescent labeling, where a  
CC protein of interest is purified and then covalently conjugated to a

CC fluorophore derivative. Novel fluorescent proteins isolated from species  
CC of the Class Anthozoa can be used as markers for gene expression and  
CC protein localization studies, and in fluorescence resonance energy  
CC transfer (FRET) reactions. They may have improved properties and better  
CC suitability for larger excitations compared to prior art fluorescent  
CC proteins such as green fluorescent protein

XX Sequence 225 AA;

Query Match 91.0%; Score 1105; DB 3; Length 225;  
Best Local Similarity 90.2%; Pred. No. 1.9e-115;  
Matches 203; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSSKNVKEKFRFVKMEGTNGHEFEIEGEGRPYEGHNTVKLKTGKGPLPFPAMD 60  
DB 1 MSCSKNVKEKFRFVKMEGTNGHEFEIEGEGRPYEGHNTVKLKTGKGPLPFPADI 60  
QY 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
DB 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
QY 121 KYKFTGVNPPSDGPNVQKKTMEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 EYKFTGVNPPSDGPNVQKKTMEASTERTLYPRDGLKGDIMALKLBGGGHYLVFEKSI 180  
QY 181 YNAKKPVQLPGYYVDSKLDITSHNEDYITVEQYERTBGRHHLFL 225  
DB 181 YNAKKPVQLPGYYVDSKLDITSHNEDYITVEQYERTBGRHHLFL 225

RESULT 96  
AAE28927  
ID AAE28927 standard; protein; 225 AA.

XX AAE28927;  
XX 27-DEC-2002 (first entry)

DE Discosoma sp. drfp583/dmfp592 (6/90) mutant protein (K15Q+K83W+T217S).  
XX  
XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
XX fluorescence activated cell sorting application; fluorescent timer;  
XX biosensor; fluorescence resonance energy transfer application; FRET;  
XX colouring agent; recombinant DNA application; analyte detection assay;  
XX sunscreen; second messenger detector; drfp583 protein; dmfp592 protein;  
XX mutant; mutein.

XX Discosoma sp.  
XX Synthetic.

OS Location/Qualifiers

XX Key  
XX MISC-difference 15 /note= "Wild-type Lys substituted with Gln"  
XX MISC-difference 83 /note= "Wild-type Lys substituted with Met"  
XX MISC-difference 217 /note= "Wild-type Thr substituted with Ser"

XX WO200268459-A2.

XX 06-SEP-2002.

XX 20-FEB-2002; 2002WO-US005749.

XX 21-FEB-2001; 2001US-0270983P.  
XX 04-DEC-2001; 2001US-00006922.

XX (CLON-) CLONTECH LAB INC.

XX Lukyanov S, Lukyanov K, Yanushkevich Y, Savitsky A, Fradkov A;

XX WPI; 2002-691654/74.

PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.

PS Disclosure: Page; 80pp; English.

XX The invention relates to nucleic acid molecules encoding non-aggregating  
XX chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
XX useful in analyte detection assays, as colouring agents, as markers in  
XX recombinant DNA applications, as sunscreens or filters, in fluorescence  
XX resonance energy transfer (FRET) applications, as biosensors in  
XX prokaryotic and eukaryotic cells, in screening assays, as second  
XX messenger detectors, in fluorescence activated cell sorting applications,  
XX in protease cleavage assays or as fluorescent timers. The present  
XX sequence is Discosoma sp. drfp583/dmfp592 (6/90) mutant protein of the  
XX invention. Note: This sequence is not shown in the specification, but is  
XX derived from Discosoma sp. drfp583/dmfp592 (6/90) wild-type protein shown  
XX as SEQ ID NO:12 (AAE28835) in page 72-73 of the specification

XX Sequence 225 AA;

Query Match 91.0%; Score 1105; DB 5; Length 225;  
Best Local Similarity 90.2%; Pred. No. 1.9e-115;  
Matches 203; Conservative 14; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSSKNVKEKFRFVKMEGTNGHEFEIEGEGRPYEGHNTVKLKTGKGPLPFPAMD 60  
DB 1 MSCSKNVKEKFRFVKMEGTNGHEFEIEGEGRPYEGHNTVKLKTGKGPLPFPADI 60  
QY 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
DB 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFDGCVTVTODSSLQDGCFTY 120  
QY 121 KYKFTGVNPPSDGPNVQKKTMEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 EYKFTGVNPPSDGPNVQKKTMEASTERTLYPRDGLKGDIMALKLBGGGHYLVFEKSI 180  
QY 181 YNAKKPVQLPGYYVDSKLDITSHNEDYITVEQYERTBGRHHLFL 225  
DB 181 YNAKKPVQLPGYYVDSKLDITSHNEDYITVEQYERTBGRHHLFL 225

RESULT 97  
AA92894  
ID AA92894 standard; protein; 230 AA.

XX AA92894;

XX 12-SEP-2000 (first entry)

XX Discosoma sp. "magenta" fluorescent protein dmfp592.

XX dmfp592; fluorescent protein; non-bioluminescent; label; marker;  
XX fluorescence resonance energy transfer reaction; FRET.

XX Discosoma sp.

XX WO200034324-A1.

XX 15-JUN-2000.

XX 10-DEC-1999; 99WO-US029412.

XX 11-DEC-1998; 98US-00210330.  
XX 19-NOV-1999; 99US-00444341.

XX (CLON-) CLONTECH LAB INC.

XX Lukyanov SA, Fradkov AF, Labas VA, Matz MV;

XX WPI; 2000-423379/36.

XX N-PSDB; AAA28958.

XX Novel fluorescent protein from non-bioluminescent *Discosoma* species,  
PT useful for fluorescent labeling and as markers.  
PS Claim 20; Page 55-56; 60pp; English.  
XX This is a novel fluorescent protein (dmpP592) from *Discosoma* sp.  
CC "magenta", a non-bioluminescent species of the genera *Anthozoa*. The  
CC fluorescent colour is observed under usual white light or ultraviolet  
CC light. It has an absorbance maximum at 573 nm and an emission maximum at  
CC 593 nm. The polynucleotides of the invention can be used to produce the  
CC proteins recombinantly, and as a source of primers and probes for  
CC identifying related proteins. The fluorescent proteins have applications  
CC in fluorescent labeling, as fluorescent markers for gene expression and  
CC protein localization studies, and in fluorescence resonance energy  
CC transfer (FRET) reactions. They may have improved properties and better  
CC suitability for larger excitations, compared to prior art fluorescent  
CC proteins such as green fluorescent protein from *Aequorea victoria*, which  
CC has an excitation maximum at 395 nm, a second excitation peak at 475 nm  
CC and an emission maximum at 510 nm.  
SQ Sequence 230 AA;  
Query Match 89.4%; Score 1085.5; DB 3; Length 230;  
Best Local Similarity 88.1%; Pred. No. 3.1e-113;  
Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;  
QY 1 MRSSKNVIKEPRFKVMEGTVNGHEFEIIEGEGRPYEGHNTVKLYTKGGPLPFAWDI 60  
DB 1 MSCSKNVIKEPRFKVMEGTVNGHEFEIIEGEGRPYEGHNTVKLYTKGGPLPFAWDI 60  
QY 61 LSPFOYGSKYVYVHPADIPDYKKLSPEEGFKMERVNNFEDGGVYVTTODSSLDDGCFIY 120  
DB 61 LSPFOYGSKYVYVHPADIPDYKKLSPEEGFKMERVNNFEDGGVYVTTODSSLDDGCFIY 120  
QY 121 KVKFTIGVNPFDGVMQKKTWGEASTERTLYPRDGVYKGEIHKALKLKDGGHYLVFPKSI 180  
DB 121 EVKFTIGVNPFDGVMQKKTWGEASTERTLYPRDGVYKGEIHKALKLKDGGHYLVFPKSI 180  
QY 181 YMAKRP-VQLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHHLFL 225  
DB 181 YMAKRP-VQLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHHPFI 226  
RESULT 98  
ID AAB01625 standard; protein; 230 AA.  
AC AAB01625;  
XX 12-DEC-2000 (first entry)  
DT 12-DEC-2000 (first entry)  
XX *Discosoma* sp. magenta fluorescent protein dmpP592.  
DE *Discosoma* sp. magenta fluorescent protein dmpP592.  
XX *Anthozoa*; fluorescent protein; fluorescence labeling; amFP486; cFP484;  
KW zFP506; zFP538; dFP583; derFP483; asFP600; dgFP512; dmpP592.  
XX *Discosoma* sp.  
OS *Discosoma* sp.  
XX WO200034526-A1.  
PN 15-JUN-2000.  
PD 15-JUN-2000.  
XX 10-DEC-1999; 99WO-US029405.  
PF 10-DEC-1999; 99WO-US029405.  
XX 11-DEC-1998; 98US-00210330.  
PR 11-DEC-1998; 98US-00210330.  
XX (CLON-) CLONTECH LAB INC.  
XX Lubyanyoy SA, Fradkov AF, Labas YA, Maltz MV;  
FI WPI; 2000-423451/36.  
XX

PT Novel method for identifying a DNA sequence encoding fluorescent proteins  
PT from non-bioluminescent *Anthozoa* which are useful for fluorescent  
PT labeling and as markers.  
PS Claim 3; Page 71-72; 73pp; English.  
XX The present sequence is *Discosoma* sp. magenta fluorescent protein  
CC dmpP592. It was isolated using the *Aequorea victoria* green fluorescent  
CC protein (GFP) sequence, which was used to design PCR primers which might  
CC isolate other fluorescent proteins from a number of species of *Anthozoa*.  
CC These were *Anemonia majano*, *Clavularia* sp., *Zoanthus* sp., *Discosoma* sp.,  
CC and *Anemonia sulcata*. The cDNA obtained was then screened in the search  
CC for sequences encoding fluorescent proteins. The other proteins found in  
CC this manner were cFP484, zFP506, zFP538, dFP583, derFP483, asFP600,  
CC dgFP512 and amFP486. These proteins can be used as fluorescent labels  
CC (for gene expression and protein localization studies and in fluorescence  
CC resonance energy transfer (FRET) studies) in place of fluorophore  
CC derivatives and luciferases, as these involve laborious processes and the  
CC latter require cofactors. They can also be used in place of GFP, which is  
CC too stable to be useful when studying short-term or repetitive events  
XX  
SQ Sequence 230 AA;  
Query Match 89.4%; Score 1085.5; DB 3; Length 230;  
Best Local Similarity 88.1%; Pred. No. 3.1e-113;  
Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;  
QY 1 MRSSKNVIKEPRFKVMEGTVNGHEFEIIEGEGRPYEGHNTVKLYTKGGPLPFAWDI 60  
DB 1 MSCSKNVIKEPRFKVMEGTVNGHEFEIIEGEGRPYEGHNTVKLYTKGGPLPFAWDI 60  
QY 61 LSPFOYGSKYVYVHPADIPDYKKLSPEEGFKMERVNNFEDGGVYVTTODSSLDDGCFIY 120  
DB 61 LSPFOYGSKYVYVHPADIPDYKKLSPEEGFKMERVNNFEDGGVYVTTODSSLDDGCFIY 120  
QY 121 KVKFTIGVNPFDGVMQKKTWGEASTERTLYPRDGVYKGEIHKALKLKDGGHYLVFPKSI 180  
DB 121 EVKFTIGVNPFDGVMQKKTWGEASTERTLYPRDGVYKGEIHKALKLKDGGHYLVFPKSI 180  
QY 181 YMAKRP-VQLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHHLFL 225  
DB 181 YMAKRP-VQLPGYVYVDSKLDITSHNEDYTVVEQYERTEGRHHPFI 226  
RESULT 99  
ID AAE00378 standard; protein; 230 AA.  
AC AAE00378;  
XX 11-SEP-2003 (revised)  
DT 11-SEP-2003 (revised)  
XX 19-JUN-2001 (first entry)  
DE *Discosoma* sp. magenta chromo/fluorescent protein, dmpP592 (NFP-9).  
XX *Anthozoa*; Chromoprotein; fluorescent protein; dmpP592; NFP-9; sunscreen;  
KW analyte detection assay; selectable marker; recombinant DNA application;  
KW biosensor; pH indicator; invivo marker; selective filter.  
XX *Discosoma* sp. magenta.  
OS *Discosoma* sp.  
XX WO200127150-A2.  
PN 19-APR-2001.  
PD 19-APR-2001.  
XX 13-OCT-2000; 2000WO-US028477.  
PF 13-OCT-2000; 2000WO-US028477.  
XX 14-OCT-1999; 99US-00418529.  
PR 14-OCT-1999; 99US-00418529.  
XX 15-OCT-1999; 99US-00418917.  
PR 15-OCT-1999; 99US-00418922.  
XX 19-NOV-1999; 99US-00444338.  
PR 19-NOV-1999; 99US-00444338.  
XX 09-DEC-1999; 99US-00457556.  
PR 09-DEC-1999; 99US-00457556.



PR 09-DEC-1999; 99US-00457898.  
 PR 09-DEC-1999; 99US-00458144.  
 PR 09-DEC-1999; 99US-00458477.  
 PR 10-DEC-1999; 99WO-US029405.  
 PR 14-JUN-2000; 2000US-0211607P.  
 PR 14-JUN-2000; 2000US-0211609P.  
 PR 14-JUN-2000; 2000US-0211626P.  
 PR 14-JUN-2000; 2000US-0211627P.  
 PR 14-JUN-2000; 2000US-0211687P.  
 PR 14-JUN-2000; 2000US-0211766P.  
 PR 14-JUN-2000; 2000US-0211880P.  
 PR 14-JUN-2000; 2000US-0211889P.  
 PR 14-JUN-2000; 2000US-0212070P.  
 (CLON-) CLONTECH LAB INC.  
 Lukyanov SA, Pradkov AP, Labas YA, Matz MV, Terékikh A;  
 WPI; 2001-266409/27.  
 N-PSDB; AAD03618.  
 An Anthozoa chromo- or fluorescent protein (P1) present in an environment other than its natural environment, useful as a label in analyte detection assays and as a selectable marker in recombinant DNA applications.  
 Claim 8; Fig 9; 69pp; English.  
 The present sequence is a Discosoma sp. magenta chromo/fluorescent protein, dmFP592 (NFP-9). NFP-9 is present in an environment other than its natural environment and has an absorbance maximum ranging from 375nm to 775nm and more usually from 560nm to 590nm and emission maximum ranging from 395nm to 795nm and more usually from 580 to 610nm. The chromoproteins or fluorescent proteins are useful as labels in analyte detection assays, as selectable markers in recombinant DNA applications, as biosensors in prokaryotic and eukaryotic cells e.g. as pH indicator, and as in vivo markers in animals. They are also useful in sunscreens and as selective filters. (Updated on 11-SEP-2003 to standardise OS field)  
 Sequence 230 AA;  
 Query Match 89.4%; Score 1085.5; DB 4; Length 230;  
 Best Local Similarity 88.1%; Pred. No. 3.1e-113;  
 Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;  
 QY 1 MSSSKVIVKEFNFKFRMREGTVNGHFEFIEGBEGRPYEGHNTVKLKTKGGPLPRAFDI 60  
 DB 1 MSSSKVIVKEFNFKFRMREGTVNGHFEFIEGBEGRPYEGHNTVKLKTKGGPLPRAFDI 60  
 QY 61 LSPQFOYGSKVYVKHPADIPDYKGLSPBGFKMERVWNPFDGCVVTVTODSSLQDGCFTY 120  
 DB 61 LSPQFOYGSKVYVKHPADIPDYKGLSPBGFKMERVWNPFDGCVVTVTODSSLQDGCFTY 120  
 QY 121 KYKFIGVNPSPDPVQOKTMGWEASTERLYPRDGYLKGFIHAKLKDGCHYLVEFKSI 180  
 DB 121 KYKFIGVNPSPDPVQOKTMGWEASTERLYPRDGYLKGFIHAKLKDGCHYLVEFKSI 180  
 QY 181 YNAKKP-VQLPGYVYVDSKLDITSHNEDYTVIOYETTERHLLFL 225  
 DB 181 YNAKKP-VQLPGYVYVDSKLDITSHNEDYTVIOYETTERHLLFL 225  
 RESULT 100  
 ABP70038 standard; protein; 230 AA.  
 AC ABP70038;  
 AC ABP70038;  
 DT 06-AUG-2003 (revised)  
 DT 22-JAN-2003 (first entry)  
 DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 240.

KM Colour facilitating molecule; CFM; green fluorescent protein; GFP;  
 KM Chromophore; biomatrix; transgenic animal; colouring agent;  
 KM flower industry; expression marker; reporter molecule; photon trap;  
 KM UV sink; sunscreen.  
 OS Unidentified.  
 PN WO200270703-A2.  
 PD 12-SEP-2002.  
 PR 01-MAR-2002; 2002WO-GB000928.  
 PR 02-MAR-2001; 2001US-0273227P.  
 PR 21-MAR-2001; 2001AU-00003874.  
 PR 15-OCT-2001; 2001US-0329816P.  
 PA (NUPA-) NUPARM LTD.  
 PA (UYOU) UNIV QUEENSLAND.  
 PA (JONE/) JONES E L.  
 PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;  
 PI Hoesgh-Gulberg IO, Prescott M;  
 DR WPI; 2002-740765/80.  
 XX Novel color-facilitating molecule for producing a biomatrix, has a  
 PT polyepide which alone/along with molecules imparts altered visual  
 PT characteristics to cells in the absence of excitation by extraneous non-  
 PT white light.  
 XX Example 19; Page 498; 510pp; English.  
 XX The invention relates to an isolated colour-facilitating molecule (CFM)  
 CC comprising a polypeptide which, in a cell, alone or together with one or  
 CC more other molecules imparts an altered visual characteristic to the cell  
 CC when visualised by a human eye. In the absence of excitation by extraneous  
 CC non-white light or particle emission. CFMs are useful for producing a  
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or  
 CC red coloured fleece. They are useful for producing coloured plant  
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other  
 CC uses include transducing or intensifying an image, providing additional  
 CC light for growing phototropic organisms e.g. algae and/or corals, for  
 CC light for growing phototropic organisms e.g. algae and/or corals, for  
 CC coating materials that experience UV damage e.g. plastics and car  
 CC upholstery. CFMs are useful in the flower industry, in the development of  
 CC new varieties of flowering plants. Other contemplated uses include,  
 CC expression markers, general reporter molecules, photon traps, UV sinks or  
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental  
 CC fungal species, and in fruits and vegetables to enhance their  
 CC marketability. CFMs embedded in a gel matrix improve image quality in  
 CC situations of distorted light spectra (biomatrix). The first all-protein  
 CC chromophore to be isolated was Green fluorescent protein (GFP). The  
 CC sequences given in records ABP6924-ABP70048 represent CFM related amino  
 CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 230 AA;  
 Query Match 89.4%; Score 1085.5; DB 5; Length 230;  
 Best Local Similarity 88.1%; Pred. No. 3.1e-113;  
 Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;  
 QY 1 MSSSKVIVKEFNFKFRMREGTVNGHFEFIEGBEGRPYEGHNTVKLKTKGGPLPRAFDI 60  
 DB 1 MSSSKVIVKEFNFKFRMREGTVNGHFEFIEGBEGRPYEGHNTVKLKTKGGPLPRAFDI 60  
 QY 61 LSPQFOYGSKVYVKHPADIPDYKGLSPBGFKMERVWNPFDGCVVTVTODSSLQDGCFTY 120  
 DB 61 LSPQFOYGSKVYVKHPADIPDYKGLSPBGFKMERVWNPFDGCVVTVTODSSLQDGCFTY 120  
 QY 121 KYKFIGVNPSPDPVQOKTMGWEASTERLYPRDGYLKGFIHAKLKDGCHYLVEFKSI 180  
 DB 121 KYKFIGVNPSPDPVQOKTMGWEASTERLYPRDGYLKGFIHAKLKDGCHYLVEFKSI 180

QY 181 YMAKP-VOLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
DB 181 YMAKKPSVOLPGYVYVDSKLDITSHNEDYTYVEQYERTQGRHHFPI 226  
RESULT 101  
ID ABP96651 standard; protein; 230 AA.  
AC ABP96651;  
XX 29-MAY-2003 (first entry)  
DE Red fluorescent protein (Dared) protein SEQ ID NO:40.  
XX  
XX Chromosome-based platform; artificial chromosome; eukaryotic chromosome;  
KW att site; integrase; recombinase; Aces; gene therapy; transgenic animal;  
KW platform artificial chromosome expression system.  
XX  
XX Discoosoma sp.  
OS Synthetic.  
OS WO200297059-A2.  
XX  
XX 05-DEC-2002.  
PD  
XX 30-MAY-2002; 2002MO-US017452.  
XX  
XX 30-MAY-2001; 2001US-0294758P.  
PR 21-MAR-2002; 2002US-0366891P.  
XX  
XX (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.  
PA  
PI Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E,  
PI Stewart S, Shellard J;  
XX  
XX MPI; 2003-140461/13.  
DR N-PSDB; ACC44649.  
XX  
XX Novel eukaryotic chromosome comprising one or many att sites which  
PT permits site-directed integration in the presence of lambda-integrase,  
PT useful for site-specific recombination-directed integration of DNA of  
PT interest.  
XX  
XX Example 1; Page 192-193; 272pp; English.  
PS  
XX The present invention describes a eukaryotic chromosome (I) comprising  
CC one or several att sites, where an att site is heterologous to the  
CC chromosome, and permits site-directed integration in the presence of  
CC lambda-integrase. Also described: (1) a platform sites that participate  
CC in recombination catalysed recombination; and (2) a method (M1) for  
CC introducing a heterologous nucleic acid into a platform artificial  
CC chromosome. (I) can be used in gene therapy. (M1) is useful for  
CC introducing a heterologous nucleic acid molecule into a platform  
CC artificial chromosome, preferably an Aces. (II) is useful for producing a  
CC transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or  
CC mammal) by introducing (II) by cell fusion, lipid-mediated transfection,  
CC by a carrier system, microinjection, microcell fusion, electroporation,  
CC microprojectile bombardment or direct DNA transfer into an embryonic  
CC cell, preferably a stem cell or an embryo. (II) comprises a heterologous  
CC nucleic acid that encodes a therapeutic product which is useful for  
CC making a library of Aces comprising random portions of a genome. ACC44612  
CC to ACC44732 and ABP96650 to ABP96657 represent sequences used in the  
CC exemplification of the present invention  
XX  
XX Sequence 230 AA;  
SQ

DB 1 MSCSKNVIKEFRRFKVMEGTVNGHFEIEGEGGPRYEGHNTVGLKVTKGGLPFAMD1 60  
QY 61 LSPQFQYGSKYVYKHGADIPDYKKLSFPEGFKWERVMANFEDGVVYTTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVYKHGADIPDYKKLSFPEGFKWERVMANFEDGVVYTTQDSSLQDGCFTY 120  
QY 121 KYKFTGVNFPSPDGVNMGKMTGMEASTBRLYPBGVYKGEIHKALKDKDGHVYVEFKSI 180  
DB 121 EVKFTGVNFPSPDGVNMGKMTGMEASTBRLYPBGVYKGEIHKALKDKDGHVYVEFKSI 180  
QY 181 YMAKP-VOLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
DB 181 YMAKKPSVOLPGYVYVDSKLDITSHNEDYTYVEQYERTQGRHHFPI 226  
RESULT 102  
ID AAE00375  
AC AAE00375 standard; protein; 205 AA.  
XX  
XX AAE00375;  
AC  
XX 11-SEP-2003 (revised)  
DT 19-JUN-2001 (first entry)  
XX  
DE Discoosoma sp. red chromo/fluorescent protein, drfp583 (NFP-6).  
XX  
XX Anthozoa; Chromoprotein; fluorescent protein; drfp583; NFP-6; sunscreen;  
KW analyte detection assay; selectable marker; recombinant DNA application;  
KW biosensor; pH indicator; invivo marker; selective filter.  
XX  
XX Discoosoma sp; red.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 150..151  
FT /note= "Encoded by TTGTATCCTCGTGGATGGCGTGTGGAAGGAGAGATT  
FT CATAGAGCTCGAAGCTGAAAGACGTCGT"  
XX  
XX WO200127150-A2.  
PD 19-APR-2001.  
XX  
XX 13-OCT-2000; 2000MO-US028477.  
XX  
XX 14-OCT-1999; 99US-00418529.  
PR 15-OCT-1999; 99US-00418917.  
PR 15-OCT-1999; 99US-00418922.  
PR 19-NOV-1999; 99US-00444338.  
PR 19-NOV-1999; 99US-00444341.  
PR 09-DEC-1999; 99US-00457556.  
PR 09-DEC-1999; 99US-00457598.  
PR 09-DEC-1999; 99US-00458144.  
PR 09-DEC-1999; 99US-00458477.  
PR 10-DEC-1999; 99MO-US029405.  
PR 14-JUN-2000; 2000US-0211607P.  
PR 14-JUN-2000; 2000US-0211626P.  
PR 14-JUN-2000; 2000US-0211627P.  
PR 14-JUN-2000; 2000US-0211627P.  
PR 14-JUN-2000; 2000US-0211687P.  
PR 14-JUN-2000; 2000US-0211766P.  
PR 14-JUN-2000; 2000US-0211880P.  
PR 14-JUN-2000; 2000US-0211888P.  
PR 14-JUN-2000; 2000US-0212070P.  
XX  
XX (CLON-) CLONTECH LAB INC.  
PA  
XX  
XX Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Teresikh A;  
PI  
XX MPI; 2001-266409/27.  
DR N-PSDB; AAD03615.  
XX  
PT An Anthozoa chromo- or fluorescent protein (P1) present in an environment  
other than its natural environment, useful as a label in analyte

PT detection assays and as a selectable marker in recombinant DNA applications.

PS Claim 8; Fig 6; 69pp; English.

XX The present sequence is a discosoma sp. red chromo/fluorescent protein, CC drp583 (NFP-6). NFP-6 is present in an environment other than its CC natural environment and has an absorbance maximum ranging from 250nm to CC 750nm and more usually from 540nm to 580nm and emission maximum ranging CC from 275nm to 775nm and more usually from 565 to 605nm. The CC chromoproteins or fluorescent proteins are useful as labels in analyte CC detection assays, as selectable markers in recombinant DNA applications, CC as biosensors in prokaryotic and eukaryotic cells e.g. as pH indicator and as in vivo markers in animals. They are also useful in sunscreens and CC as selective filters. Note: The specification also refers to following CC mutants of this sequence: (a) E5 with substitution V105A and S197T, (b) CC E5down with substitution S197T, (c) E57 with substitutions V105A, I161, CC S197A, (d) AG4 with substitutions V71W, V105A, S197T and (e) AG4H with CC substitutions V71W, V105A, Y120H and S197T. However, the positions of S CC in (a)-(e) and I in (c) do not match with this wild type sequence.

CC (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 205 AA:

QY Query Match 87.1%; Score 1058; DB 4; Length 205; Best Local Similarity 88.4%; Pred. No. 3.3e-110; Matches 199; Conservative 2; Mismatches 4; Indels 20; Gaps 1;

DB 1 MSSSKNVIKFPMFKFRMGVTGVNGHFEI EGEGEGRPYEGHNTVKLKATKGGPLPRAWDI 60

QY 1 MSSSKNVIKFPMFKFRMGVTGVNGHFEI EGEGEGRPYEGHNTVKLKATKGGPLPRAWDI 60

DB 61 LSPQFGYSKVVYKHPADIPDYKLSFPBGFKMERVMNFEDGCVTVTQDSSIQDCCFIT 120

QY 61 LSPQFGYSKVVYKHPADIPDYKLSFPBGFKMERVMNFEDGCVTVTQDSSIQDCCFIT 120

DB 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGVLKGEIHKALKIKDGGHYLVEFKSI 180

QY 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGVLKGEIHKALKIKDGGHYLVEFKSI 180

DB 121 KSSSLALNPPSDGPVWQKTMGWEASTERL-----GHYLVFEKSI 160

QY 181 YMAKKEVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

DB 161 YMAKKEVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 205

RESULT 103

AAE00384 ID AAE00384 standard; protein; 205 AA.

XX AAE00384;

XX 19-JUN-2001 (first entry)

XX Discosoma sp. red drp583 (NFP-6) mutant E5up (V105A).

XX Anthozoa; Chromoprotein; fluorescent protein; drp583; NFP-6; sunscreen; KM analyte detection assay; selectable marker; recombinant DNA application; KW biosensor; pH indicator; in vivo marker; selective filter; mutant; murein.

XX Discosoma sp; red.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 105 /note= "wild type Val substituted with Ala"

XX MO200127150-A2.

XX 19-APR-2001.

XX 13-OCT-2000; 2000MO-US028477.

XX 14-OCT-1999; 99US-00418529.

PR 15-OCT-1999; 99US-00418917.

PR 15-OCT-1999; 99US-00418922.

PR 19-NOV-1999; 99US-00444338.

PR 19-NOV-1999; 99US-00444341.

PR 09-DEC-1999; 99US-00457556.

PR 09-DEC-1999; 99US-00457898.

PR 09-DEC-1999; 99US-00458144.

PR 09-DEC-1999; 99US-00458477.

PR 10-DEC-1999; 99MO-US029405.

PR 14-JUN-2000; 2000US-0211607P.

PR 14-JUN-2000; 2000US-0211609P.

PR 14-JUN-2000; 2000US-0211626P.

PR 14-JUN-2000; 2000US-0211627P.

PR 14-JUN-2000; 2000US-0211687P.

PR 14-JUN-2000; 2000US-0211766P.

PR 14-JUN-2000; 2000US-0211800P.

PR 14-JUN-2000; 2000US-0211888P.

PR 14-JUN-2000; 2000US-0212070P.

XX (CLON-) CLONTECH LAB INC.

PA Lukyanov SA, Fradkov AF, Labas YA, Maltz MV, Terzikikh A;

PI MPI; 2001-266409/27.

XX An Anthozoa chromo- or fluorescent protein (P1) present in an environment PT other than its natural environment, useful as a label in analyte PT detection assays and as a selectable marker in recombinant DNA applications.

XX Disclosure; Page; 69pp; English.

XX The present sequence is a discosoma sp. red chromo/fluorescent protein, CC drp583 (NFP-6) mutant E5up (V105A). NFP-6 is present in an environment CC other than its natural environment and has an absorbance maximum ranging CC from 250nm to 750nm and more usually from 540nm to 580nm and emission CC maximum ranging from 275nm to 775nm and more usually from 565 to 605nm. CC The chromoproteins or fluorescent proteins are useful as labels in CC analyte detection assays, as selectable markers in recombinant DNA CC applications, as biosensors in prokaryotic and eukaryotic cells e.g. as CC pH indicator and as in vivo markers in animals. They are also useful in CC sunscreens and as selective filters. Note: The present sequence is not CC shown in the specification, but derived from drp583 (NFP-6) sequence CC (AAE00375) shown in figure 6

XX Sequence 205 AA:

QY Query Match 86.8%; Score 1054; DB 4; Length 205; Best Local Similarity 88.0%; Pred. No. 9.3e-110; Matches 198; Conservative 2; Mismatches 5; Indels 20; Gaps 1;

DB 1 MSSSKNVIKFPMFKFRMGVTGVNGHFEI EGEGEGRPYEGHNTVKLKATKGGPLPRAWDI 60

QY 1 MSSSKNVIKFPMFKFRMGVTGVNGHFEI EGEGEGRPYEGHNTVKLKATKGGPLPRAWDI 60

DB 61 LSPQFGYSKVVYKHPADIPDYKLSFPBGFKMERVMNFEDGCVTVTQDSSIQDCCFIT 120

QY 61 LSPQFGYSKVVYKHPADIPDYKLSFPBGFKMERVMNFEDGCVTVTQDSSIQDCCFIT 120

DB 61 LSPQFGYSKVVYKHPADIPDYKLSFPBGFKMERVMNFEDGCVTVTQDSSIQDCCFIT 120

QY 121 KYKFIGVNPSPDGPVWQKTMGWEASTERLYPRDGVLKGEIHKALKIKDGGHYLVEFKSI 180

DB 121 KSSSLALNPPSDGPVWQKTMGWEASTERL-----GHYLVFEKSI 160

QY 181 YMAKKEVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

DB 161 YMAKKEVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 205

RESULT 104

AAE00383 ID AAE00383 standard; protein; 205 AA.

XX AAE00383;

XX 19-JUN-2001 (first entry)  
DT  
XX Diacosoma sp. red drfp583 (NFP-6) mutant E8 (N42H).  
DE  
XX Anthozoa; Chromoprotein; fluorescent protein; drfp583; NFP-6; sunscreen;  
KW analyte detection assay; selectable marker; recombinant DNA application;  
KW biosensor; pH indicator; in vivo marker; selective filter; mutant; muteln.  
XX  
OS Diacosoma sp; red.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 42 /note= "Wild type Asn substituted with His"  
FT  
XX WO200127150-A2.  
XX  
XX 19-APR-2001.  
PD  
XX 13-OCT-2000; 2000MO-US028477.  
PF  
XX 14-OCT-1999; 99US-00418529.  
XX 15-OCT-1999; 99US-00418917.  
XX 15-OCT-1999; 99US-00418922.  
XX 19-NOV-1999; 99US-00444338.  
XX 19-NOV-1999; 99US-00444341.  
XX 09-DEC-1999; 99US-00457556.  
XX 09-DEC-1999; 99US-00457898.  
XX 09-DEC-1999; 99US-00458144.  
XX 09-DEC-1999; 99US-00458477.  
XX 10-DEC-1999; 99MO-US029405.  
XX 14-JUN-2000; 2000US-0211607P.  
XX 14-JUN-2000; 2000US-0211609P.  
XX 14-JUN-2000; 2000US-0211626P.  
XX 14-JUN-2000; 2000US-0211627P.  
XX 14-JUN-2000; 2000US-0211687P.  
XX 14-JUN-2000; 2000US-0211766P.  
XX 14-JUN-2000; 2000US-0211880P.  
XX 14-JUN-2000; 2000US-0211888P.  
XX 14-JUN-2000; 2000US-0212070P.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
XX Lukyanov SA, Fradkov AP, Labas YA, Matz MV, Terexikh A;  
XX WPI, 2001-266409/27.  
DR  
XX An Anthozoa chromo- or fluorescent protein (Pl) present in an environment  
XX other than its natural environment, useful as a label in analyte  
XX detection assays and as a selectable marker in recombinant DNA  
XX applications.  
XX  
XX Disclosure; Page; 69pp; English.  
XX  
XX The present sequence is a Diacosoma sp. red chromo/fluorescent protein,  
XX drfp583 (NFP-6) mutant E8 (N42H). NFP-6 is present in an environment  
XX other than its natural environment and has an absorbance maximum ranging  
XX from 250nm to 750nm and more usually from 540nm to 580nm and emission  
XX maximum ranging from 275nm to 775nm and more usually from 565 to 605nm.  
XX The chromoproteins or fluorescent proteins are useful as labels in  
XX analyte detection assays, as selectable markers in recombinant DNA  
XX applications, as biosensors in prokaryotic and eukaryotic cells e.g. as  
XX pH indicator and as in vivo markers in animals. They are also useful in  
XX sunscreens and as selective filters. Note: The present sequence is not  
XX shown in the specification, but derived from drfp583 (NFP-6) sequence  
XX (AAE00375) shown in figure 6  
XX  
XX Sequence 205 AA;  
SQ

QY 1 MRSSKNVKEFRMRKRVMEGTVNGHEPEIEGEGGRPYEGHNTVTKTKGKGPILPEANDI 60  
DB 1 MRSSKNVKEFRMRKRVMEGTVNGHEPEIEGEGGRPYEGHNTVTKTKGKGPILPEANDI 60  
QY 61 LSPFOYGSKYVGHKPADIPYKLLSPFEGFKMERVMNFEDGVTVTVQDSLDGCFIY 120  
DB 61 LSPFOYGSKYVGHKPADIPYKLLSPFEGFKMERVMNFEDGVTVTVQDSLDGCFIY 120  
QY 121 KVFPIGVNPSDGPVMOKKTMGWEASTERLYPRDGLKGEIHKALKKODGGHYLVEPKSI 180  
DB 121 KSSSLALNFPDGPVMOKKTMGWEASTERL-----GHYLVFPKSI 160  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYEOYERTEGRHHLFL 225  
DB 161 YMAKKPVQLPGYVYVDSKLDITSHNEDYTYEOYERTEGRHHLFL 205  
RESULT 105  
ADCC24133  
ID ADC24133 standard; protein; 225 AA.  
XX  
XX ADC24133;  
AC  
XX 18-DEC-2003 (first entry)  
DT  
XX  
XX Diacosoma red fluorescent protein variant mRFP1.  
XX  
XX Diacosoma red fluorescent protein; Dared; AB interface; AC interface;  
KW fluorescent protein variant; transcription induction detection;  
KW fluorescence energy resonance transfer; FRRET; protein kinase;  
KW protein phosphatase; ion indicator; mutant; mRFP1; muteln.  
XX  
XX Synthetic.  
OS  
OS Diacosoma.  
XX  
XX US2003059835-A1.  
XX  
XX 27-MAR-2003.  
PD  
XX  
XX 10-APR-2002; 2002US-00121258.  
XX  
XX 26-FEB-2001; 2001US-00794308.  
XX 24-MAY-2001; 2001US-00866538.  
XX  
XX (TSIE//) TSIE N Y.  
XX (CAMP//) CAMPBELL R E.  
XX  
XX Tsien RY, Campbell RB;  
PI  
XX WPI, 2003-743764/70.  
XX  
XX N-PSDB; ADC24132.  
DR  
XX  
XX Novel polynucleotide sequence encoding Diacosoma red fluorescent protein  
XX variant having a reduced propensity to oligomerize, useful for detecting  
XX transcriptional activity.  
XX  
XX Disclosure; SEQ ID NO 8; 67pp; English.  
XX  
XX The invention describes a polynucleotide sequence (I) encoding a  
XX Diacosoma red fluorescent protein (Dared) variant having a reduced  
XX propensity to oligomerize, comprising amino acid substitutions at the AB  
XX and/or AC interfaces of the wild-type Dared sequence (SI) comprising 225  
XX amino acids, given in the specification, where the substitutions result  
XX in reduced propensity of the Dared variant to form tetramers. (I) is  
XX useful for detecting transcriptional activity by providing a host cells  
XX containing a vector which comprises (I) operatively linked to an  
XX expression control sequence, and an unit to assay the variant fluorescent  
XX protein fluorescence, and assaying fluorescence of the variant  
XX fluorescent protein produced by (VII), where variant fluorescent protein  
XX fluorescence is indicative of transcriptional activity. A polynucleotide  
XX encoding a fusion protein is useful for the analysis of in vivo  
XX localisation or trafficking of a polypeptide of interest. A polypeptide

CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcription, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as  $Ca^{2+}$ ,  $Zn^{2+}$ , for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (1) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This is the amino acid sequence of  
CC Discosoma red fluorescent protein variant mRFP1.

SQ Sequence 225 AA;

Query Match 84.1%; Score 1021; DB 7; Length 225;

Best Local Similarity 86.9%; Pred. No. 5.6e-106;  
Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

```
QY 1 MRSKNVIEFMRFKYRMGGTNGHFEIIEGSEGRPYEGHNTVTKYTKGGPLPPAMDI 60
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 1 MASSEDEVIEKFMKFKYRMGGSVNGHFEIIEGSEGRPYEGTQYAKLKYTKGGPLPPAMDI 60
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 LSPQFYGSKAVYKHPADIPDYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGEFITY 120
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 61 LSPQFYGSKAVYKHPADIPDYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGEFITY 120
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 121 KYKFLIGVNFPSDGPVWQKKTWGEASTERLYPRDGVYKGEIHKALYKDGHYLVFESKI 180
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 121 KYKFLIGVNFPSDGPVWQKKTWGEASTERLYPRDGVYKGEIHKALYKDGHYLVFESKI 180
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 181 YNAKKPVQLPGYTYVDSKLDITSHNEDYTVTEOYERTEGRH 221
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 181 YNAKKPVQLPGAYKTDIKLDITSHNEDYTVTEOYERABGRH 221
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 106

ADL46210 ADL46210 standard; protein; 225 AA.

XX AC ADL46210;

DT 20-MAY-2004 (first entry)

DE Discosoma red fluorescent protein (DsRed) variant mRFP1 protein.

XX red fluorescent protein; DsRed; fluorescence; red wavelength;

KW oligomerization; tetramerization; immunoassay; hybridization assay.

OS Discosoma sp.

PN MO2003086446-A1.

XX 23-OCT-2003.

XX 09-APR-2003; 2003WC-US010879.

XX 10-APR-2002; 2002US-00121258.

XX 29-JUL-2002; 2002US-00209208.

XX (REGC ) UNIV CALIFORNIA.

XX Tsien RY, Campbell RE, Baird GS;

XX WPI; 2003-845265/78.

XX N-PSDB; ADL46211.

XX New monomeric and dimeric Anthozoan fluorescent protein variants with  
XX reduced propensity to oligomerize, and encoding polynucleotides, useful  
XX in molecular biology, e.g. in immunoassays or in tracking protein

PT movement in cells.

XX Disclosure; SEQ ID NO 8; 166pp; English.

PS The invention relates to a polynucleotide sequence encoding a Discosoma  
XX red fluorescent protein (DsRed) variant having a reduced propensity to  
XX oligomerize. The protein variant comprises one or more amino acid  
XX substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
XX sequence, where the substitutions result in reduced propensity of the  
XX DsRed variant to form tetramers and where the variant displays detectable  
XX fluorescence of at least one red wavelength. The composition and methods  
XX are useful in producing red fluorescent proteins having reduced  
XX propensity for oligomerization, especially tetramerization. The protein  
XX may be used in molecular biology and in other scientific applications,  
XX such as in immunoassays or hybridization assays, or in tracking the  
XX movement of proteins in cells. This sequence corresponds to the DsRed  
XX variant mRFP1 protein.

SQ Sequence 225 AA;

Query Match 84.1%; Score 1021; DB 7; Length 225;

Best Local Similarity 86.9%; Pred. No. 5.6e-106;  
Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

```
QY 1 MRSKNVIEFMRFKYRMGGTNGHFEIIEGSEGRPYEGHNTVTKYTKGGPLPPAMDI 60
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 1 MASSEDEVIEKFMKFKYRMGGSVNGHFEIIEGSEGRPYEGTQYAKLKYTKGGPLPPAMDI 60
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 LSPQFYGSKAVYKHPADIPDYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGEFITY 120
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 61 LSPQFYGSKAVYKHPADIPDYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGEFITY 120
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 121 KYKFLIGVNFPSDGPVWQKKTWGEASTERLYPRDGVYKGEIHKALYKDGHYLVFESKI 180
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 121 KYKFLIGVNFPSDGPVWQKKTWGEASTERLYPRDGVYKGEIHKALYKDGHYLVFESKI 180
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 181 YNAKKPVQLPGYTYVDSKLDITSHNEDYTVTEOYERTEGRH 221
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 181 YNAKKPVQLPGAYKTDIKLDITSHNEDYTVTEOYERABGRH 221
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 107

ADQ59566 ADQ59566 standard; protein; 225 AA.

XX AC ADQ59566;

DT 07-OCT-2004 (first entry)

DE Discosoma monomerised red fluorescent protein mRFP1.

XX optically detectable signal; protein fragment complementation assay; PCA;

KW molecular interaction detection; biomolecular interaction detection;

KW high-throughput screening; high-content screening; pathway mapping;

KW drug mechanism-of-action study; biosensor; diagnostic;

XX red fluorescent protein; spectral shift; mRFP1.

XX Discosoma.

XX US2004137528-A1.

XX 15-JUL-2004.

XX 01-DEC-2003; 2003US-00724178.

XX 02-FEB-1998; 98US-00017412.

XX 07-FEB-2000; 2000US-00499464.

XX 24-MAY-2002; 2002US-00154758.

XX 29-JAN-2003; 2003US-00353090.

XX 09-APR-2003; 2003US-0461133P.

XX (MICH/) WATSON MICHNICK S W.

PA (MACD/) MACDONALD M. L.  
 BA (LAME/) LAMERDIN J.  
 XX Watson Michnick SW, Macdonald ML, Lamerdin J;  
 PT MPI; 2004-533363/51.  
 DR N-PSDB; ADQ59565.  
 CC Composition useful in protein fragment complementation assays for drug  
 PT discovery and high-throughput screening, comprising complementary  
 PT fragments of protein or mutant protein, generating optically detectable  
 PT signal when associated.  
 PS Disclosure: SEQ ID NO 16; 34pp; English.  
 CC The invention describes a composition (1) comprising complementary  
 CC fragments of a protein or mutant protein, where the fragments generate an  
 CC optically detectable signal when associated, and each of the mutant  
 CC protein fragments is fused to a separate molecule. Also described are:  
 CC protein fragment complementation assays (PCAs) for detection of molecular  
 CC interactions, involving reassembling separate fragments from an optically  
 CC detectable protein, and detecting the reassembly by units of  
 CC reconstruction of activity of the optically detectable protein, where the  
 CC reassembly of the fragments is operated by the interaction of molecular  
 CC domains fused to each fragment, and is independent of other molecular  
 CC processes; detecting biomolecular interaction, involving selecting an  
 CC appropriate optically detectable protein, effecting fragmentation of the  
 CC optically detectable protein such that the fragmentation results in  
 CC reversible loss of protein function, fusing or attaching fragments of the  
 CC reversibly detectable protein separately to other molecules, reassociating  
 CC the protein fragments through interactions of the molecules that are  
 CC fused or attached to the fragments, and detecting the resulting optical  
 CC signal; designing and engineering of PCAs based on fluorescent protein,  
 CC and a method and composition for the construction of multi-color PCAs.  
 CC (1) is useful in PCAs and other assays for drug discovery, target  
 CC validation, high-throughput screening, high-content screening, pathway  
 CC mapping, drug mechanism-of-action studies, biosensors and diagnostics.  
 CC (1) is useful for engineering different colour PCAs for a variety of  
 CC applications in biology and biotechnology. This is the amino acid  
 CC sequence of a Discosoma monomerised red fluorescent protein from which  
 CC PCA fragments can be isolated.  
 SQ Sequence 225 AA;  
 Query Match 84.1%; Score 1021; DB 8; Length 225;  
 Best Local Similarity 86.9%; Pred. No. 5.6e-106;  
 Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0;  
 Oy 1 MRSSKNVTKERFMRKVMGTVNGHGFETEGEGRPRYEGHNTVTLKVTYKGGPLPFANDI 60  
 Db 1 MASSSEDIKKEFRKFRVMEGSGVNGHGFETEGEGRPRYEGTQATLKTGKGGPLPFANDI 60  
 Oy 61 LSPFOYSGSKYVYVHPADIPDYKLLSPFEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120  
 Db 61 LSPFOYSGSKAYVHPADIPDYLLKLSPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120  
 Oy 121 KVKRIGVNPSPSGRPMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFESKI 180  
 Db 121 KVKRIGTNPSPSGRPMQKTMGWEASTERYMPEDALKGELIKMRLLKLDGGHYDAEVKTT 180  
 Oy 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
 Db 181 YMAKKPVQLPGAYKTDIKLDTSHNEDYTIIVEQYERABGRH 221  
 RESULT 108  
 ID AEA54910 standard; protein; 747 AA.  
 AC AEA54910;  
 XX 25-AUG-2005 (first entry)  
 XX

DE Fluorescence-related EGFP-caespase-3 DEVD linker-mRFP(x2) fusion protein.  
 XX fluorescence; protein localization; cellular transport; organelle;  
 KW mitochondria; fusion protein; enhanced green fluorescent protein; EGFP;  
 KW caespase-3.  
 XX  
 OS Unidentified.  
 OS Chimeric.  
 XX  
 FN W02005054464-A1.  
 XX  
 PD 16-JUN-2005.  
 XX  
 PE 03-DEC-2004; 2004WO-JP018437.  
 XX  
 PR 03-DEC-2003; 2003JP-00404472.  
 XX  
 PR 27-JAN-2004; 2004JP-00018344.  
 XX  
 PA (RIKE ) RIKEN KK.  
 XX (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.  
 XX Miyawaki A, Kogure T, Hama H, Kinjo M, Saito K, Karasawa S;  
 PT Aizaki T;  
 DR MPI; 2005-444979/45.  
 XX N-PSDB; AEA54911.  
 CC Novel fluorescent protein existing as monomer and isolated from Fungia  
 PT and Montipora species, useful for labeling and analyzing location and  
 PT movement of intracellular organelles.  
 PS Example 10; SEQ ID NO 49; 218pp; Japanese.  
 CC The invention relates to a novel fluorescent protein existing as a  
 CC monomer and comprising any one of SEQ ID NO. 1-29 (odd SEQ ID numbers)  
 CC given in the specification, or any one of the sequences chosen from SEQ  
 CC ID NO. 1-29 (odd SEQ ID numbers), with one or more amino acid  
 CC substitutions, deletions and/or additions, and having fluorescent  
 CC characteristics. The fusion protein of the invention may be useful for  
 CC analyzing the location or movement of a protein in a cell in vivo and for  
 CC labeling and analyzing the location or movement of intracellular  
 CC organelles, such as mitochondria. The current sequence is that of the  
 CC fluorescence-related enhanced green fluorescent protein (EGFP)-caespase-3  
 CC DEVD linker-mRFP(x2) fusion protein of the invention.  
 SQ Sequence 747 AA;  
 Query Match 84.1%; Score 1021; DB 9; Length 747;  
 Best Local Similarity 86.9%; Pred. No. 3.3e-105;  
 Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0;  
 Oy 1 MRSSKNVTKERFMRKVMGTVNGHGFETEGEGRPRYEGHNTVTLKVTYKGGPLPFANDI 60  
 Db 269 MASSSEDIKKEFRKFRVMEGSGVNGHGFETEGEGRPRYEGTQATLKTGKGGPLPFANDI 328  
 Oy 61 LSPFOYSGSKYVYVHPADIPDYKLLSPFEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120  
 Db 329 LSPFOYSGSKAYVHPADIPDYLLKLSPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 388  
 Oy 121 KVKRIGVNPSPSGRPMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFESKI 180  
 Db 389 KVKRIGTNPSPSGRPMQKTMGWEASTERYMPEDGALKGELIKMRLLKLDGGHYDAEVKTT 448  
 Oy 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
 Db 449 YMAKKPVQLPGAYKTDIKLDTSHNEDYTIIVEQYERABGRH 489  
 RESULT 109  
 ID ADL46281 standard; protein; 225 AA.  
 AC ADL46281;  
 XX  
 XX

XX 20-MAY-2004 (first entry)  
DT  
XX  
DE Diacosoma red fluorescent protein (Dered) variant mRFP1.1 coding sequence.  
XX  
XX red fluorescent protein; Dered; fluorescence; red wavelength;  
KM oligomerization; tetramerization; immunoassay; hybridization assay.  
XX  
XX Diacosoma sp.  
OS Synthetic.  
XX  
XX WO2003086446-A1.  
PN  
XX  
PD 23-OCT-2003.  
XX  
XX 09-APR-2003; 2003WO-US010879.  
PF  
XX 10-APR-2002; 2002US-00121258.  
PR 29-JUL-2002; 2002US-00209208.  
PR  
XX  
XX (REGC ) UNIV CALIFORNIA.  
PA  
XX  
XX Tsien RY, Campbell RE, Baird GS;  
PI  
XX  
XX WPI; 2003-845265/78.  
DR  
XX  
XX New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.  
PT  
XX  
XX Claim 70; SEQ ID NO 79, 166pp; English.  
PS  
XX  
XX The invention relates to a polynucleotide sequence encoding a Diacosoma  
CC red fluorescent protein (Dered) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type Dered  
CC sequence, where the substitutions result in reduced propensity of the  
CC Dered variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the Dered  
CC variant mRFP1.1 protein which contains a Q66M and a T147S mutation.  
XX  
XX  
XX Sequence 225 AA;  
XQ

	Query Match	83.4%	Score 1012	DA 7	Length 225
	Best Local Similarity	86.0%	Pred. No. 5.8e-105		
	Matches 190	Conservative 6	Mismatches 25	Indels 0	Gaps 0
Qy	1	MRSSKNVTKFEMRFRVMEGTVNGHFEPIEGEGEGRPYEGHNTVTKLTKTKGGLPFPAADI	60		
Db	1	MASSSEDVKEFMRFRVMEGTVNGHFEPIEGEGEGRPYEGGTOTALTKTKGGLPFPAADI	60		
Qy	61	LSPOQVYSKYYVGHKPADIPDYKKLSFPEGPFMEWVNMPEQGGVTVYQDSSLQDGGCTY	120		
Db	61	LSPOQMYSSKIVKHPADIPDYLLKSPFEGPFMEWVNMPEQGGVTVYQDSSLQDGGCTY	120		
Qy	121	KVKKFGVNFPSDGPVMOQKTMGWEASTERYLPERDVLKGEIHKALKDQGHYIVFESKI	180		
Db	121	KVKLRGTNFPSDGPVMOQKTMGWEASTERYLPEDALKEIIMRLKLDQGHYDAEVKTT	180		
Qy	181	YMAKPPVQLPGYYVDSLDTISNHEDTYIVQYERTGRH	221		
Db	181	YMAKPPVQLPGAYKTDIKLDTISNHEDTYIVQYERABGRH	221		

Search completed: January 11, 2006, 02:08:28  
Job time : 81 secs

**THIS PAGE BLANK (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 11, 2006, 02:05:28 ; Search time 16 Seconds  
(without alignments)  
1353.048 Million cell updates/sec

Title: US-10-006-922A-12

Perfect score: 1214  
Sequence: 1 MRSSKNVTKERMPKVMKG.....EDYTIYQYERTGSRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%

Listing first 500 summaries

Database : PIR\_80:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

No matches found

Search completed: January 11, 2006, 02:10:16

Job time : 16 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using sw model

Run on: January 11, 2006, 02:04:58 ; Search time 73 Seconds  
(without alignments)  
2174.574 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFKRMFKVRMEG.....EDYIVEQYERTGRHFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Database: UniProt\_05.80:\*  
1: uniprot\_prot:\*  
2: uniprot\_trembl:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	1 RFP DISSP	Q9uY8 discosoma s
2	1196	98.5	236	2 Q66ND7_9CNID	Q66nd7 discosoma s
3	1193	98.3	225	2 Q6KPB5_9CNID	Q6Kpb5 discosoma s
4	1191	98.1	236	2 Q66ND8_9CNID	Q66nd8 discosoma s
5	1085.5	89.4	230	2 Q9GJ7_9CNID	Q9Gj7 discosoma s
6	985	81.1	226	2 Q5S3G8_9CNID	Q5S3g8 discosoma s
7	972	80.1	226	2 Q5S3G7_9CNID	Q5S3g7 discosoma s

ALIGNMENTS

RESULT 1  
RFP DISSP STANDARD; PRT; 225 AA.  
AC Q9U6Y8;  
DT 01-FEB-2005 (Rel. 46, Created)  
DT 01-FEB-2005 (Rel. 46, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
OS Red fluorescent protein drfp583 (DisRed).  
OS discosoma sp. (See anemone).  
OC Eukaryota; Metazoa; Chordata; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomacidae; Discosoma.  
OX NCBI\_TaxID=86600;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=99436614; Pubmed=10504696; DOI=10.1038/13657;  
RA Matz M.V., Pradlov A.P., Labas Y.A., Savitsky A.P., Zaratsky A.G.,  
RA Markelov M.L., Lukyanov S.A.,  
RT "Fluorescent proteins from nonbioluminescent Anthozoa species";  
RL Nat. Biotechnol. 17:969-973(1999).

[2]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX PubMed=1101896; DOI=10.1038/81992;  
RA Wall M.A., Socolich M., Ranganathan R.,  
RT "The structural basis for red fluorescence in the tetrameric GFP homolog DisRed";  
RL Nat. Struct. Biol. 7:1133-1138(2000).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND SUBUNIT.  
RX PubMed=11209050; DOI=10.1073/pnas.98.2.462;  
RA Yarbrough D., Wachter R.M., Kallio K., Matz M.V., Remington S.J.,  
RT "Refined crystal structure of DisRed, a red fluorescent protein from coral, at 2.0-A resolution";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:462-467(2001).  
CC -1- FUNCTION: Thought to play a role in photoprotection of the coral's resident symbiotic microalgae's photosystems from photoinhibition caused by high light levels found near the surface of coral reefs. In deeper water, the fluorescence may be to convert blue light into longer wavelengths more suitable for use in photosynthesis by the microalgal symbionts.  
CC -1- BIOPHYSICOCHEMICAL PROPERTIES:  
Absorption:  
Abs(max)=558 nm;  
Note=Exhibits a smaller absorbance peak at 494 nm. The broad fluorescence emission spectrum peaks at 583 nm;  
CC -1- SUBUNIT: Homotrimer.  
CC -1- PTM: Contains a chromophore consisting of modified amino acid residues. The chromophore is formed by autocatalytic backbone condensation between Xaa-N and Gly-(N+2), oxidation of Tyr-(N+1) to dihydroxytyrosine, and formation of a double bond to the alpha-amino nitrogen of residue Xaa-N. Maturation of the chromophore requires nothing other than molecular oxygen.  
CC -1- BIOTECHNOLOGY: Fluorescent proteins have become a useful and ubiquitous tool for making chimeric proteins, where they function as a fluorescent protein tag. Typically they tolerate N- and C-terminal fusion to a broad variety of proteins. They have been expressed in most known cell types and are used as a noninvasive fluorescent marker in living cells and organisms. They enable a wide range of applications where they have functioned as a cell lineage tracer, reporter of gene expression, or as a measure of protein-protein interactions.  
CC -1- SIMILARITY: Belongs to the GFP family.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
CC EMBL, AF168419; AA03369.1; -, mRNA.  
CC PDB, 1G7K; X-ray; A/B/C/D=2-225.  
CC PDB, 1GGX; X-ray; A/B/C/D=1-225.  
CC InterPro: IPR011584; GFP related.  
CC InterPro: IPR000786; Green\_fl\_protein.  
CC Pfam: PF01353; GFP; 1.  
CC PRINTS: PR01229; GFP/DOESCENT.  
CC DR ProDom; PD013756; Green\_fl\_protein; 1.  
CC 3D-structure; Chromophore; Luminescence; Photoprotein.  
CC MOD RES 67 67 (2)-2,3-didehydrotyrosine.  
CC FT CROSSLINK 66 68 2-iminomethyl-5-imidazolone (Gln-Gly).  
CC SEQUENCE 225 AA; 25931 MW; FBFPA5369778F68 CRC64;  
Query Match 100.0%; Score 1214; DB 1; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-96;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFKRMFKVRMEGTVNGHFEIYEGSGEGRPYGGHNVVKAKTKGGLPFWAMD 60  
DB 1 MRSSKNVKEFKRMFKVRMEGTVNGHFEIYEGSGEGRPYGGHNVVKAKTKGGLPFWAMD 60  
QY 61 LSPQFGYSKVYVKHADIIPDYKLSFPPGFKMERVMNFDGQVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFGYSKVYVKHADIIPDYKLSFPPGFKMERVMNFDGQVTVTQDSSLQDGCFTY 120

```
OY 121 KVFIFGVNPPSDGPMQKKTGMWEASTERYLPDGLKGEIHKALKLKDGGHYLVEPKSI 180
DB 121 KVFIFGVNPPSDGPMQKKTGMWEASTERYLPDGLKGEIHKALKLKDGGHYLVEPKSI 180
OY 181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTYVEQYERTGRRHHLFL 225
DB 181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTYVEQYERTGRRHHLFL 225

RESULT 2
O66ND7_9CNID
ID Q66ND7_9CNID PRELIMINARY; PRT; 236 AA.
AC Q66ND7;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DB Enhanced red fluorescent protein R+.
OS Discosoma sp. RC-2004.
OC Eukaryota; Metazoa; Chndaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
RN NCB1_TaxID=289055;
RN NUCLEOTIDE SEQUENCE.
RA Carter R.W., Gibbs P.D.L., Schmale M.C.;
RT "Cloning of Anthozoan Fluorescent Protein Genes.";
RL Submitted (JUL-2004) to the EMBL/Genbank/DBJ databases.
DB EMBL; AY679107; AAU0444.1; -, mRNA.
GO GO:0006091; P:generation of precursor metabolites and energy; IEA.
DR InterPro; IPR000786; GFP related.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PRO1229; GFPLORESCENT.
DR PRODOM; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 236 AA; 27032 MW; BB3844BDE06829EFO CRC64;

Query Match 98.5%; Score 1196; DB 2; Length 236;
Best Local Similarity 98.2%; Pred. No. 3.9e-95;
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRSSKNVYKEPMRFKVMMEGVNNGHEFEIEBEGSGRPYEGHNTVTKLVTKGGLPFPANDI 60
DB 1 MSCSKNVYKEPMRFKVMMEGVNNGHEFEIEBEGSGRPYEGHNTVTKLVTKGGLPFPANDI 60
OY 61 LSPQFGYSKYVVGHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120
DB 61 LSPQFGYSKYVVGHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120
OY 121 KVFIFGVNPPSDGPMQKKTGMWEASTERYLPDGLKGEIHKALKLKDGGHYLVEPKSI 180
DB 121 KVFIFGVNPPSDGPMQKKTGMWEASTERYLPDGLKGEIHKALKLKDGGHYLVEPKSI 180
OY 181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTYVEQYERTGRRHHLFL 225
DB 181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTYVEQYERTGRRHHLFL 225

RESULT 3
O66F85_9CNID
ID Q66F85_9CNID PRELIMINARY; PRT; 225 AA.
AC Q66F85;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DB Orange fluorescent protein FP586.
OS Discosoma sp. JW-2002.
OC Eukaryota; Metazoa; Chndaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
RN NCB1_TaxID=208461;
RN NUCLEOTIDE SEQUENCE.
RA Wiedemann J., Girod A.;
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
```

```
DR EMBL; AF545828; AAQ11987.1; -, mRNA.
DR HSSP; P42312; 189C.
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.
DR InterPro; IPR011584; GFP-related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PRO1229; GFPLORESCENT.
DR PRODOM; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 225 AA; 25751 MW; E151D0E497AA23FA CRC64;

Query Match 99.3%; Score 1193; DB 2; Length 225;
Best Local Similarity 98.2%; Pred. No. 6.7e-95;
Matches 221; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRSSKNVYKEPMRFKVMMEGVNNGHEFEIEBEGSGRPYEGHNTVTKLVTKGGLPFPANDI 60
DB 1 MSCSKNVYKEPMRFKVMMEGVNNGHEFEIEBEGSGRPYEGHNTVTKLVTKGGLPFPANDI 60
OY 61 LSPQFGYSKYVVGHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120
DB 61 LSPQFGYSKYVVGHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120
OY 121 KVFIFGVNPPSDGPMQKKTGMWEASTERYLPDGLKGEIHKALKLKDGGHYLVEPKSI 180
DB 121 KVFIFGVNPPSDGPMQKKTGMWEASTERYLPDGLKGEIHKALKLKDGGHYLVEPKSI 180
OY 181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTYVEQYERTGRRHHLFL 225
DB 181 YMAKKPVQLPGYVVYDSKLDITSHNEDYTYVEQYERTGRRHHLFL 225

RESULT 4
O66ND8_9CNID
ID Q66ND8_9CNID PRELIMINARY; PRT; 236 AA.
AC Q66ND8;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DB Red fluorescent protein R1.
OS Discosoma sp. RC-2004.
OC Eukaryota; Metazoa; Chndaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
RN NCB1_TaxID=289055;
RN NUCLEOTIDE SEQUENCE.
RA Carter R.W., Gibbs P.D.L., Schmale M.C.;
RT "Cloning of Anthozoan Fluorescent Protein Genes.";
RL Submitted (JUL-2004) to the EMBL/Genbank/DBJ databases.
DB EMBL; AY679106; AAU0444.1; -, mRNA.
GO GO:0006091; P:generation of precursor metabolites and energy; IEA.
DR InterPro; IPR011584; GFP related.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PRO1229; GFPLORESCENT.
DR PRODOM; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 236 AA; 27042 MW; BB3844BCD6834EF3 CRC64;

Query Match 98.1%; Score 1191; DB 2; Length 236;
Best Local Similarity 97.8%; Pred. No. 1.1e-94;
Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRSSKNVYKEPMRFKVMMEGVNNGHEFEIEBEGSGRPYEGHNTVTKLVTKGGLPFPANDI 60
DB 1 MSCSKNVYKEPMRFKVMMEGVNNGHEFEIEBEGSGRPYEGHNTVTKLVTKGGLPFPANDI 60
OY 61 LSPQFGYSKYVVGHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120
DB 61 LSPQFGYSKYVVGHPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120
OY 121 KVFIFGVNPPSDGPMQKKTGMWEASTERYLPDGLKGEIHKALKLKDGGHYLVEPKSI 180
DB 121 KVFIFGVNPPSDGPMQKKTGMWEASTERYLPDGLKGEIHKALKLKDGGHYLVEPKSI 180
```

QY 181 YNAKRPVQLPGYVYVDSKLDITSHNEDYTYVEQYRTGRHLLFL 225  
 DB 181 YNAKRPVQLPGYVYVDSKLDITSHNEDYTYVEQYRTGRHLLFL 225

RESULT 5  
 Q9GTU7\_9CNID PRELIMINARY; PRT; 230 AA.  
 AC Q9GTU7;  
 DT 01-MAR-2003 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Red fluorescent protein.  
 GN Name=FP93;  
 OS Discosoma sp. SSAL-2000.  
 OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
 OC Discosomatidae; Discosoma.  
 NCBI\_TaxId=137428;  
 RX NUCLEOTIDE SEQUENCE.  
 RP MEDLINE=20434599; PubMed=10981720; DOI=10.1016/S0014-5793(00)01895-0;  
 RA Fradkov A.F., Chen Y., Ding L., Barsova E.V., Matz M.V.,  
 RA Lukyanov S.A.;  
 RT "Novel fluorescent protein from Discosoma coral and its mutants  
 RT possesses a unique far-red fluorescence.";  
 RL FEBS Lett. 479:127-130(2000).  
 DR EMBL; AF272711; AAG16224.1; -, mRNA.  
 DR HSSP; 09U6Y8; 1GKX.  
 DR GO; GO:0006091; P.generation of precursor metabolites and energy; IEA.  
 DR InterPro; IPR011584; GFP\_related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.  
 DR PRINTS; PR01229; GFPLORESCENT.  
 DR SEQUENCE 230 AA; 26370 MW; 5215B1B436D67E51 CRC64;

Query Match 89.4%; Score 1085.5; DB 2; Length 230;  
 Best Local Similarity 88.1%; Pred. No. 1.3e-85;  
 Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;

QY 1 MSSSKNVIEKPMRFRKYMGEVTVNGHFEIIEGEGRPRYEGHNTVKKLTGKGPLPRAWDI 60  
 DB 1 MCSCKNVIEKPMRFRKYMGEVTVNGHFEIIEGEGRPRYEGHNTVKKLTGKGPLPRAWDI 60

QY 61 LSPQFQYGSKVYVHPADIPDYKLSFPGGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
 DB 61 LSPQFQYGSKVYVHPADIPDYKLSFPGGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120

QY 121 KVFVIGVNFPSDGPVWQKKTWGEASTERLYPRDGVLYKGEIHKALKKDGCHYLVFKSI 180  
 DB 121 KVFVIGVNFPSDGPVWQKKTWGEASTERLYPRDGVLYKGEIHKALKKDGCHYLVFKSI 180

QY 121 EVKFIGVNFPSDGPVWQKKTWGEASTERLYPRDGVLYKGEIHKALKKDGCHYLVFKSI 180  
 DB 121 EVKFIGVNFPSDGPVWQKKTWGEASTERLYPRDGVLYKGEIHKALKKDGCHYLVFKSI 180

QY 181 YNAKRPVQLPGYVYVDSKLDITSHNEDYTYVEQYRTGRHLLFL 225  
 DB 181 YNAKRPVQLPGYVYVDSKLDITSHNEDYTYVEQYRTGRHLLFL 225

RESULT 6  
 Q5S3G8\_9CNID PRELIMINARY; PRT; 226 AA.  
 AC Q5S3G8;  
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Fluorescent protein raspberry.  
 OS Discosoma sp. LM-2004.  
 OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
 OC Discosomatidae; Discosoma.  
 NCBI\_TaxId=301246;  
 RX NUCLEOTIDE SEQUENCE.  
 RP PubMed=15556995; DOI=10.1073/pnas.0407752101;  
 RA Wang L., Jackson W.C., Steinbach P.A., Tsien R.Y.;  
 RT "Evolution of new nonantibody proteins via iterative somatic

RT hypermutation.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:16745-16749(2004).  
 DR EMBL; AY786536; AAV65486.1; -, mRNA.  
 DR GO; GO:0006091; P.generation of precursor metabolites and energy; IEA.  
 DR InterPro; IPR011584; GFP\_related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.  
 DR PRINTS; PR01229; GFPLORESCENT.  
 DR ProDom; PD013756; Green\_fl\_protein; 1.  
 DR SEQUENCE 226 AA; 25513 MW; 159ABFB2D453A5B CRC64;

Query Match 81.1%; Score 985; DB 2; Length 226;  
 Best Local Similarity 85.3%; Pred. No. 6.3e-77;  
 Matches 185; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

QY 5 KNVIEKPMRFRKYMGEVTVNGHFEIIEGEGRPRYEGHNTVKKLTGKGPLPRAWDIISPO 64  
 DB 6 BEVIEKPMRFRKYMGEVTVNGHFEIIEGEGRPRYEGHNTVKKLTGKGPLPRAWDIISPO 65

QY 65 FOYGSKVYVHPADIPDYKLSFPGGFKMERVMNPFEDGVTVTQDSSLQDGCFTYKVKF 124  
 DB 66 CMYGSKGVYVHPADIPDYKLSFPGGFKMERVMNPFEDGVTVTQDSSLQDGCFTYKVKL 125

QY 125 IGVNFPDGPVWQKKTWGEASTERLYPRDGVLYKGEIHKALKKDGCHYLVFKSIYNAK 184  
 DB 126 KRTNFPDGPVWQKKTWGEASTERLYPRDGVLYKGEIHKALKKDGCHYLVFKSIYNAK 185

QY 185 KPVQLPGVYVDSKLDITSHNEDYTYVEQYRTGRH 221  
 DB 186 KPVQLPGVYVDSKLDITSHNEDYTYVEQYRTGRH 222

RESULT 7  
 Q5S3G7\_9CNID PRELIMINARY; PRT; 226 AA.  
 AC Q5S3G7;  
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Fluorescent protein plum.  
 OS Discosoma sp. LM-2004.  
 OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
 OC Discosomatidae; Discosoma.  
 NCBI\_TaxId=301246;  
 RX NUCLEOTIDE SEQUENCE.  
 RP PubMed=15556995; DOI=10.1073/pnas.0407752101;  
 RA Wang L., Jackson W.C., Steinbach P.A., Tsien R.Y.;  
 RT "Evolution of new nonantibody proteins via iterative somatic  
 RT hypermutation.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:16745-16749(2004).  
 DR EMBL; AY786537; AAV65487.1; -, mRNA.  
 DR SMR; Q5S3G7; 8-222.  
 DR GO; GO:0006091; P.generation of precursor metabolites and energy; IEA.  
 DR InterPro; IPR011584; GFP\_related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.  
 DR PRINTS; PR01229; GFPLORESCENT.  
 DR ProDom; PD013756; Green\_fl\_protein; 1.  
 DR SEQUENCE 226 AA; 25590 MW; 6DF38CAD2AB28BED CRC64;

Query Match 80.1%; Score 972; DB 2; Length 226;  
 Best Local Similarity 83.4%; Pred. No. 8.4e-76;  
 Matches 181; Conservative 8; Mismatches 28; Indels 0; Gaps 0;

QY 5 KNVIEKPMRFRKYMGEVTVNGHFEIIEGEGRPRYEGHNTVKKLTGKGPLPRAWDIISPO 64  
 DB 6 BEVIEKPMRFRKYMGEVTVNGHFEIIEGEGRPRYEGHNTVKKLTGKGPLPRAWDIISPO 65

QY 65 FOYGSKVYVHPADIPDYKLSFPGGFKMERVMNPFEDGVTVTQDSSLQDGCFTYKVKF 124  
 DB 66 IMVGSKVYVHPADIPDYKLSFPGGFKMERVMNPFEDGVTVTQDSSLQDGCFTYKVKV 125

QY 125 IGVNFPDGPVWQKKTGMEASTERLYPRDGVKGEIHKALKLDKGHYLVEFKSIYMAK 184  
 DB 126 RGTNFPDGPVWQKKTGMEASERMPEDGALKGEMGRLRLKDGCHYDAEVKTTYMAK 185  
 QY 185 KPVQLPGYTYVDSKUDITSHNEDYTIIVEQYERTEGRH 221  
 DB 186 KPVQLPGAYKTDIKLDITSHNEDYTIIVEQYERABEGRH 222

Search completed: January 11, 2006, 02:09:49  
 Job time : 74 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 02:06:44 ; Search time 23 Seconds

(without alignments)  
808.784 Million cell updates/sec

Title: US-10-006-922a-12

Perfect score: 1214

Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYIVQEYERTGRHHLFL 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%

Maximum Match 100%

Listing first 500 summaries

Database:

Issued Patents, AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5-COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/6-COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/7-COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/8-COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/9-COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1214	100.0	225	2	US-09-866-538-12	Sequence 12, Appl
2	1214	100.0	225	2	US-09-865-291-12	Sequence 12, Appl
3	1210	99.7	240	2	US-10-152-296-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-866-538-12  
Sequence 12, Application US/09866538  
Patent No. 6852849  
GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSIEN, Roger  
APPLICANT: Campbell, Robert  
TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
FILE REFERENCE: REGEN1530-2  
CURRENT APPLICATION NUMBER: US/09/866,538  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: Patent version 3.0  
SEQ ID NO 12  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
US-09-866-538-12

Query Match 100.0%; Score 1214; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 3e-132;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRSSKNVKEFMRFKVRMEGTVNGHEFEI	EGEGGRPYEGHNTVYKLTGKGPPLPFAMDI	60
DB	1	MRSSKNVKEFMRFKVRMEGTVNGHEFEI	EGEGGRPYEGHNTVYKLTGKGPPLPFAMDI	60
QY	61	LSPOFGYSKYVVKHPADIPDYKLSFPEGFK	MEVWVFEDGVVTVTQDSSLQDGCFTY	120
DB	61	LSPOFGYSKYVVKHPADIPDYKLSFPEGFK	MEVWVFEDGVVTVTQDSSLQDGCFTY	120
QY	121	KYKFIGVNPSPDGPVWQKKTWGEASTERLY	PRDGVLTGKEIHKALKDGGHYLVEFKSI	180
DB	121	KYKFIGVNPSPDGPVWQKKTWGEASTERLY	PRDGVLTGKEIHKALKDGGHYLVEFKSI	180
QY	181	YMAKRPVQLPGYVYVDSKLDITSHNEDYTI	VEQYERTGRHHLFL 225	
DB	181	YMAKRPVQLPGYVYVDSKLDITSHNEDYTI	VEQYERTGRHHLFL 225	

#### RESULT 2

US-09-865-291-12  
Sequence 12, Application US/09865291  
Patent No. 6900304  
GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSIEN, Roger  
APPLICANT: TING, Alice  
APPLICANT: ZHANG, Jin  
TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION  
FILE REFERENCE: REGEN1550  
CURRENT APPLICATION NUMBER: US/09/865,291  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: Patent version 3.0  
SEQ ID NO 12  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
US-09-865-291-12

Query Match 100.0%; Score 1214; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 3e-132;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRSSKNVKEFMRFKVRMEGTVNGHEFEI	EGEGGRPYEGHNTVYKLTGKGPPLPFAMDI	60
DB	1	MRSSKNVKEFMRFKVRMEGTVNGHEFEI	EGEGGRPYEGHNTVYKLTGKGPPLPFAMDI	60
QY	61	LSPOFGYSKYVVKHPADIPDYKLSFPEGFK	MEVWVFEDGVVTVTQDSSLQDGCFTY	120
DB	61	LSPOFGYSKYVVKHPADIPDYKLSFPEGFK	MEVWVFEDGVVTVTQDSSLQDGCFTY	120
QY	121	KYKFIGVNPSPDGPVWQKKTWGEASTERLY	PRDGVLTGKEIHKALKDGGHYLVEFKSI	180
DB	121	KYKFIGVNPSPDGPVWQKKTWGEASTERLY	PRDGVLTGKEIHKALKDGGHYLVEFKSI	180
QY	181	YMAKRPVQLPGYVYVDSKLDITSHNEDYTI	VEQYERTGRHHLFL 225	
DB	181	YMAKRPVQLPGYVYVDSKLDITSHNEDYTI	VEQYERTGRHHLFL 225	

#### RESULT 3

US-10-152-296-2  
Sequence 2, Application US/10152296  
Patent No. 6723537  
GENERAL INFORMATION:  
APPLICANT: Peelle, Beau  
APPLICANT: Rigel Pharmaceuticals, Incorporated  
TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
FILE REFERENCE: 021044-000110US

```
; CURRENT APPLICATION NUMBER: US/10/152,296
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/291,871
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mammalian
; OTHER INFORMATION: codon-optimized variant (DsRED) of Discosoma sp.
; OTHER INFORMATION: "red" red fluorescent protein (RFP)
US-10-152-296-2

Query Match          99.7%; Score 1210; DB 2; Length 240;
Best Local Similarity 99.6%; Pred. No. 9.5e-132;
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVTKKEMRFKVRMEGVTVNGHFEPIEGRBGRPYEGHNTVTKLTKTKGGPLPFANDI 60
        : |||||
Db       2 VRSSKNVTKKEMRFKVRMEGVTVNGHFEPIEGRBGRPYEGHNTVTKLTKTKGGPLPFANDI 61

QY      61 LSPQFGSKVYVKHPADIPIYKKLSPEPGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120
        : |||||
Db       62 LSPQFGSKVYVKHPADIPIYKKLSPEPGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 121

QY      121 KVKFIGVNFPSDGPVMQKTMGMWMASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180
        : |||||
Db       122 KVKFIGVNFPSDGPVMQKTMGMWMASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 181

QY      181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIYEQYERTGRRHHLFL 225
        : |||||
Db       182 YMAKRPVQLPGYYVDSKLDITSHNEDYTIYEQYERTGRRHHLFL 226
```

Search completed: January 11, 2006, 02:10:50  
Job time : 23 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OK protein - protein search, using sw model

Run on: January 11, 2006, 02:07:04 / Search time 62 Seconds  
(without alignments)  
1516.316 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFRFRKVRMEG.....EDYIVQEYRTERGRHFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 75

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Database : Published Applications AA Main:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	1214	100.0	225	US-09-999-745-67	Sequence 67, Appl
2	1214	100.0	225	US-09-866-538-12	Sequence 12, Appl
3	1214	100.0	225	US-09-794-308-12	Sequence 12, Appl
4	1214	100.0	225	US-09-865-291-12	Sequence 12, Appl
5	1214	100.0	225	US-10-006-922-12	Sequence 12, Appl
6	1214	100.0	225	US-10-081-864-8	Sequence 8, Appl
7	1214	100.0	225	US-10-121-258-1	Sequence 1, Appl
8	1214	100.0	225	US-10-315-920-2	Sequence 2, Appl
9	1214	100.0	225	US-10-132-067-4	Sequence 4, Appl
10	1214	100.0	225	US-10-370-570-56	Sequence 56, Appl
11	1214	100.0	225	US-10-406-618-32	Sequence 32, Appl
12	1214	100.0	225	US-10-433-640-13	Sequence 13, Appl
13	1214	100.0	225	US-10-311-030-7	Sequence 7, Appl
14	1214	100.0	225	US-10-845-484-3	Sequence 3, Appl
15	1214	100.0	225	US-10-885-988-12	Sequence 12, Appl
16	1214	100.0	225	US-10-656-029-22	Sequence 22, Appl
17	1214	100.0	225	US-10-857-622-12	Sequence 12, Appl
18	1214	100.0	225	US-10-505-486-26	Sequence 26, Appl
19	1214	100.0	225	US-10-844-064A-2	Sequence 2, Appl
20	1214	100.0	225	US-10-931-304-1	Sequence 1, Appl
21	1214	100.0	487	US-10-343-977-1	Sequence 1, Appl
22	1214	100.0	487	US-10-343-977-1	Sequence 1, Appl
23	1214	100.0	506	US-10-343-977-2	Sequence 2, Appl
24	1214	100.0	506	US-10-343-977-2	Sequence 2, Appl
25	1214	100.0	547	US-10-343-977-3	Sequence 3, Appl
26	1214	100.0	547	US-10-343-977-3	Sequence 3, Appl
27	1210	99.7	226	US-10-311-030-9	Sequence 9, Appl

28	1210	99.7	240	US-10-152-296-2	Sequence 2, Appl
29	1210	99.7	240	US-10-739-656-2	Sequence 2, Appl
30	1210	99.7	545	US-10-214-932-52	Sequence 52, Appl
31	1210	99.7	548	US-10-214-932-76	Sequence 76, Appl
32	1207	99.4	225	US-10-121-258-20	Sequence 20, Appl
33	1207	99.4	225	US-10-315-920-4	Sequence 4, Appl
34	1207	99.4	225	US-10-931-304-20	Sequence 20, Appl
35	1196	98.5	236	US-10-314-936-2	Sequence 2, Appl
36	1196	98.5	236	US-10-314-936-4	Sequence 4, Appl
37	1196	98.5	236	US-11-021-014-2	Sequence 2, Appl
38	1196	98.5	236	US-11-021-014-4	Sequence 4, Appl
39	1195.5	98.5	226	US-10-724-178-12	Sequence 12, Appl
40	1191	98.1	225	US-10-315-920-6	Sequence 6, Appl
41	1186	97.7	225	US-10-442-148A-7	Sequence 7, Appl
42	1186	97.7	229	US-10-442-148A-8	Sequence 8, Appl
43	1160	95.6	225	US-10-121-258-4	Sequence 4, Appl
44	1160	95.6	225	US-10-931-304-4	Sequence 4, Appl
45	1153	95.0	225	US-10-121-258-24	Sequence 24, Appl
46	1153	95.0	225	US-10-931-304-24	Sequence 24, Appl
47	1132	93.2	225	US-10-423-688A-41	Sequence 41, Appl
48	1121	92.3	226	US-10-121-258-6	Sequence 6, Appl
49	1121	92.3	226	US-10-931-304-6	Sequence 6, Appl
50	1121	92.3	602	US-11-052-001-6	Sequence 6, Appl
51	1119	92.2	225	US-10-006-922-44	Sequence 44, Appl
52	1119	92.2	225	US-10-081-864-12	Sequence 12, Appl
53	1095	90.2	234	US-10-931-304-81	Sequence 81, Appl
54	1095	90.2	476	US-10-931-304-106	Sequence 106, Appl
55	1085.5	89.4	230	US-10-006-922-18	Sequence 18, Appl
56	1085.5	89.4	230	US-10-161-403-40	Sequence 40, Appl
57	1085.5	89.4	230	US-11-006-076-40	Sequence 40, Appl
58	1046	86.2	205	US-10-006-922-46	Sequence 46, Appl
59	1021	84.1	225	US-10-121-258-8	Sequence 8, Appl
60	1021	84.1	225	US-10-724-178-16	Sequence 16, Appl
61	1021	84.1	225	US-10-931-304-8	Sequence 8, Appl
62	1014	83.5	225	US-10-931-304-108	Sequence 108, Appl
63	1012	83.4	225	US-10-931-304-79	Sequence 79, Appl
64	1007	82.9	225	US-10-931-304-104	Sequence 104, Appl
65	1002	82.5	225	US-10-931-304-100	Sequence 100, Appl
66	994	81.9	236	US-10-931-304-85	Sequence 85, Appl
67	992	81.7	236	US-10-931-304-89	Sequence 89, Appl
68	988	81.5	236	US-10-931-304-102	Sequence 102, Appl
69	988	81.4	236	US-10-931-304-83	Sequence 83, Appl
70	987	81.3	236	US-10-931-304-92	Sequence 92, Appl
71	986	81.2	235	US-10-931-304-94	Sequence 94, Appl
72	973	80.1	236	US-10-931-304-96	Sequence 96, Appl
73	973	80.1	236	US-10-931-304-98	Sequence 98, Appl
74	960	79.1	236	US-10-931-304-87	Sequence 87, Appl
75	848	69.9	183	US-10-724-178-1041	Sequence 1041, Ap

ALIGNMENTS

RESULT 1  
US-09-999-745-67  
Sequence 67, Application US/0999745  
Patent No. US20020157120A1  
GENERAL INFORMATION:  
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: Tsien, Roger Y.  
APPLICANT: Baird, Geoffrey  
TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS  
FILE REFERENCE: REGN1470-1  
CURRENT APPLICATION NUMBER: US/09/999,745  
CURRENT FILING DATE: 2001-10-23  
PRIOR APPLICATION NUMBER: 09/316,920  
PRIOR FILING DATE: 1999-05-21  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 67  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.

US-09-999-745-67

Query Match 100.0%; Score 1214; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVQLKVTYKGGPLPFAMD 60  
DB 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVQLKVTYKGGPLPFAMD 60  
QY 61 LSPQFQYGSKYVVKHPADIDPYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVVKHPADIDPYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGVPMQKTMGWEASTERYLRDGVLKGEIHKALKDGGHYLVEFKSI 180  
DB 121 KYKFIGVNPSPDGVPMQKTMGWEASTERYLRDGVLKGEIHKALKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
DB 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 2

US-09-866-538-12  
Sequence 12, Application US/09866538  
Publication No. US20030032088A1

GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSJEN, Roger  
APPLICANT: CAMPBELL, Robert  
TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
FILE REFERENCE: REGEN1530-2  
CURRENT APPLICATION NUMBER: US/09/866,538  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 12  
LENGTH: 225  
TYPE: PRT  
ORGANISM: *Discoosoma* sp.  
US-09-866-538-12

Query Match 100.0%; Score 1214; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVQLKVTYKGGPLPFAMD 60  
DB 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVQLKVTYKGGPLPFAMD 60  
QY 61 LSPQFQYGSKYVVKHPADIDPYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVVKHPADIDPYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGVPMQKTMGWEASTERYLRDGVLKGEIHKALKDGGHYLVEFKSI 180  
DB 121 KYKFIGVNPSPDGVPMQKTMGWEASTERYLRDGVLKGEIHKALKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
DB 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 3

US-09-794-308-12  
Sequence 12, Application US/09794308  
Publication No. US20030170911A1

GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSJEN, Roger  
APPLICANT: ZACHARIAS, David  
APPLICANT: BAIRD, Geoffrey

TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
FILE REFERENCE: REGEN1530  
CURRENT APPLICATION NUMBER: US/09/794,308  
CURRENT FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 12  
LENGTH: 225  
TYPE: PRT  
ORGANISM: *Discoosoma* sp.  
US-09-794-308-12

Query Match 100.0%; Score 1214; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVQLKVTYKGGPLPFAMD 60  
DB 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVQLKVTYKGGPLPFAMD 60  
QY 61 LSPQFQYGSKYVVKHPADIDPYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVVKHPADIDPYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGVPMQKTMGWEASTERYLRDGVLKGEIHKALKDGGHYLVEFKSI 180  
DB 121 KYKFIGVNPSPDGVPMQKTMGWEASTERYLRDGVLKGEIHKALKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
DB 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 4

US-09-865-291-12  
Sequence 12, Application US/09865291  
Publication No. US20030186229A1

GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSJEN, Roger  
APPLICANT: TING, Alice  
APPLICANT: ZHANG, Jin  
TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION  
FILE REFERENCE: REGEN1550  
CURRENT APPLICATION NUMBER: US/09/865,291  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 12  
LENGTH: 225  
TYPE: PRT  
ORGANISM: *Discoosoma* sp.  
US-09-865-291-12

Query Match 100.0%; Score 1214; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVQLKVTYKGGPLPFAMD 60  
DB 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVQLKVTYKGGPLPFAMD 60  
QY 61 LSPQFQYGSKYVVKHPADIDPYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVVKHPADIDPYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGVPMQKTMGWEASTERYLRDGVLKGEIHKALKDGGHYLVEFKSI 180  
DB 121 KYKFIGVNPSPDGVPMQKTMGWEASTERYLRDGVLKGEIHKALKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
DB 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

```

RESULT 5
US-10-006-922-12
/ Sequence 12, Application US/10006922
/ Publication No. US20020197676A1
/ GENERAL INFORMATION:
/ APPLICANT: Lukanov, Sergey A
/ APPLICANT: Lukanov, Arcady F.
/ APPLICANT: Labas, Yulii A.
/ APPLICANT: Matz, Mikhail V.
/ APPLICANT: Terekhin, Alexey
/ TITLE OF INVENTION: No. US20020197676A1 Chromophores/fluorophores and
/ TITLE OF INVENTION: Methods for using the same
/ FILE REFERENCE: CLOW-035CIP
/ CURRENT APPLICATION NUMBER: US/10/006,922
/ CURRENT FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 09/120,330
/ PRIOR FILING DATE: 1998-12-11
/ PRIOR APPLICATION NUMBER: 09/457,898
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/458,144
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/458,477
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/457,556
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/444,338
/ PRIOR FILING DATE: 1999-11-19
/ NUMBER OF SEQ ID NOS: 46
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Discosoma species
US-10-006-922-12

Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSKKNVKEFMFKYRMGEVTNGHEFEIIEGEGRPYEGHNTVTKLVTGKGPLPFAWDI 60
    |||||
DB 1 MRSKKNVKEFMFKYRMGEVTNGHEFEIIEGEGRPYEGHNTVTKLVTGKGPLPFAWDI 60

QY 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 120
    |||||
DB 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 120

QY 121 KYKFTGVNPPSDGPRVWQKKTGMGEASTERLYPRDGVTKGEIHKALKDKDGHYLVFKSI 180
    |||||
DB 121 KYKFTGVNPPSDGPRVWQKKTGMGEASTERLYPRDGVTKGEIHKALKDKDGHYLVFKSI 180

QY 121 KYKFTGVNPPSDGPRVWQKKTGMGEASTERLYPRDGVTKGEIHKALKDKDGHYLVFKSI 180
    |||||
DB 121 KYKFTGVNPPSDGPRVWQKKTGMGEASTERLYPRDGVTKGEIHKALKDKDGHYLVFKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225
    |||||
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 6
US-10-081-864-8
/ Sequence 8, Application US/10081864
/ Publication No. US20030022287A1
/ GENERAL INFORMATION:
/ APPLICANT: Lukanov, Sergey
/ APPLICANT: Lukanov, Konstantin
/ APPLICANT: Yanushevich, Yuriy
/ APPLICANT: Savitschey, Alexander
/ APPLICANT: Pradkov, Arcady
/ TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and
/ TITLE OF INVENTION: Methods for using the same
/ FILE REFERENCE: CLOW-067
/ CURRENT APPLICATION NUMBER: US/10/081,864
/ CURRENT FILING DATE: 2002-06-19

```

```

/ PRIOR APPLICATION NUMBER: 10/006,922
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 60/270,983
/ PRIOR FILING DATE: 2001-02-21
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 8
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Discosoma sp.
US-10-081-864-8

Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSKKNVKEFMFKYRMGEVTNGHEFEIIEGEGRPYEGHNTVTKLVTGKGPLPFAWDI 60
    |||||
DB 1 MRSKKNVKEFMFKYRMGEVTNGHEFEIIEGEGRPYEGHNTVTKLVTGKGPLPFAWDI 60

QY 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 120
    |||||
DB 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 120

QY 121 KYKFTGVNPPSDGPRVWQKKTGMGEASTERLYPRDGVTKGEIHKALKDKDGHYLVFKSI 180
    |||||
DB 121 KYKFTGVNPPSDGPRVWQKKTGMGEASTERLYPRDGVTKGEIHKALKDKDGHYLVFKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225
    |||||
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 7
US-10-121-258-1
/ Sequence 1, Application US/10121258
/ Publication No. US20030059835A1
/ GENERAL INFORMATION:
/ APPLICANT: Tsien, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: UC083.1CP2CP1
/ CURRENT APPLICATION NUMBER: US/10/121,258
/ CURRENT FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Discosoma sp.
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(225)
/ OTHER INFORMATION: wild-type Dared
US-10-121-258-1

Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6,6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSKKNVKEFMFKYRMGEVTNGHEFEIIEGEGRPYEGHNTVTKLVTGKGPLPFAWDI 60
    |||||
DB 1 MRSKKNVKEFMFKYRMGEVTNGHEFEIIEGEGRPYEGHNTVTKLVTGKGPLPFAWDI 60

QY 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 120
    |||||
DB 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVNNFEDGGVVTYQDSSLQDGCFTY 120

```

QY	QY
121	121
KVAFIGVNPSPDGPVWQKKTGMWMASTERLIPRDOVLKGEITHKALKLKDGSHYIVVERKSL	KVAFIGVNPSPDGPVWQKKTGMWMASTERLIPRDOVLKGEITHKALKLKDGSHYIVVERKSL
121	121
KVAFIGVNPSPDGPVWQKKTGMWMASTERLIPRDOVLKGEITHKALKLKDGSHYIVVERKSL	KVAFIGVNPSPDGPVWQKKTGMWMASTERLIPRDOVLKGEITHKALKLKDGSHYIVVERKSL
181	181
YVAKKEVPLPGYVYVYVDSKLDITSHNEDYTIYEQYERTGRRHPL	YVAKKEVPLPGYVYVYVDSKLDITSHNEDYTIYEQYERTGRRHPL
181	181
YVAKKEVPLPGYVYVYVDSKLDITSHNEDYTIYEQYERTGRRHPL	YVAKKEVPLPGYVYVYVDSKLDITSHNEDYTIYEQYERTGRRHPL

```

RESULT 8
US-10-315-920-2
/ Sequence 2, Application US/10315920
/ Publication No. US20030175809A1
/ GENERAL INFORMATION:
/ APPLICANT: Fradkov, Arcady Fedorovich
/ APPLICANT: Tersikh, Alexey
/ TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
/ TITLE OF INVENTION: FOR THEIR USE
/ FILE REFERENCE: CLON-077C1P
/ CURRENT APPLICATION NUMBER: US/10/315,920
/ CURRENT FILING DATE: 2002-12-09
/ PRIOR APPLICATION NUMBER: 60/211,607
/ PRIOR FILING DATE: 2000-06-14
/ PRIOR APPLICATION NUMBER: PCT/US01/19097
/ PRIOR FILING DATE: 2001-06-13
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Dicoboma sp.
/ US-10-315-920-2

```

```

RESULT 9
US-10-132-067-4
/ Sequence 4, Application US/10132067
/ Publication No. US20030203355A1
/ GENERAL INFORMATION:
/ APPLICANT: Bradbury, Andrew
/ APPLICANT: Zeytun, Ahmet
/ APPLICANT: Waldo, Geoffrey
/ APPLICANT: The Regents of the University of California
/ TITLE OF INVENTION: Fluorobodies: Binding Ligands with Intrinsic
/ TITLE OF INVENTION: Fluorescence
/ FILE REFERENCE: 021362-000600US
/ CURRENT APPLICATION NUMBER: US/10/132,067
/ CURRENT FILING DATE: 2002-04-24
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 4
/
/ TYPE: PRT
/ LENGTH: 225
/

```

```

; ORGANISM: Discosoma sp.
;
; FEATURE:
; OTHER INFORMATION: red fluorescent protein (dsRED)
US-10-132-067-4

```

	Query Match	100.0%	Score 1214;	DB 4;	Length 225;
	Best Local Similarity	100.0%	Pred. No. 6.6e-114;		
	Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	MRSSKNVYKEFMKFKVMEGTGVNGHFEFIEBEGSGRPYEIGNHTYKLKVTKGSPLEFPADI	60		
Db	1	MRSSKNVYKEFMKFKVMEGTGVNGHFEFIEBEGSGRPYEIGNHTYKLKVTKGSPLEFPADI	60		
QY	61	LSPOFGSKYYVGHPADIPYKLSPEEGKRMRRVMMFEGGIVYTYQDSLDGCTTY	120		
Db	61	LSPOFGSKYYVGHPADIPYKLSPEEGKRMRRVMMFEGGIVYTYQDSLDGCTTY	120		
QY	121	KVKFIGNVFPDGGVMMOKTGMGEASTERLAPRDGLKGEIHKALKLKDGGHYLVAFESI	180		
Db	121	KVKFIGNVFPDGGVMMOKTGMGEASTERLAPRDGLKGEIHKALKLKDGGHYLVAFESI	180		
QY	181	YMAKPPVQLPGYYVDSKLDITTSNEDTYTYEQRTEREGRHHTL	225		
Db	181	YMAKPPVQLPGYYVDSKLDITTSNEDTYTYEQRTEREGRHHTL	225		

```

RESULT 10
US-10-370-570-56
; Sequence 56, Application US/10370570
; Publication NO. US2003021971A1
; GENERAL INFORMATION:
; APPLICANT: DAHL, Soren Weis et al.
; TITLE OF INVENTION: FLUOROPHORE COMPLEMENTATION PRODUCTS
; FILE REFERENCE: 3759-0130P
; CURRENT APPLICATION NUMBER: US/10/370,570
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discooma sp.
US-10-370-570-56

```

RESULT 11  
US-10-406-618-32  
; Sequence 32, Application US/10406618  
; Publication No. US20030219814A1  
; GENERAL INFORMATION:  
; APPLICANT: Wan, David Chi-Chong  
; APPLICANT: Ip, Denis Ter-Ming  
; APPLICANT: The Chinese University of Hong Kong

TITLE OF INVENTION: No. US20030219814A1el Fluorescent Proteins  
FILE REFERENCE: 016285-34-1  
CURRENT APPLICATION NUMBER: US/10/406,618  
CURRENT FILING DATE: 2003-04-02  
PRIOR APPLICATION NUMBER: US 60/387,968  
PRIOR FILING DATE: 2002-06-11  
PRIOR APPLICATION NUMBER: US 60/370,598  
PRIOR FILING DATE: 2002-04-05  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 32  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Spontaneously  
OTHER INFORMATION: Fluorescent protein Dered.  
US-10-406-618-32

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVMREGTVNGHGFIEGSGRPRYGNTTVKLTGKGPLPPAMD 60  
1 MSSKNVKEFMRFKVMREGTVNGHGFIEGSGRPRYGNTTVKLTGKGPLPPAMD 60  
DB 1 LSPQFGSKVYVKKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120  
61 LSPQFGSKVYVKKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
DB 121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
QY 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
DB 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 12  
US-10-433-640-13  
Sequence 13, Application US/10433640  
Publication No. US20040115792A1  
GENERAL INFORMATION:  
APPLICANT: Lichenberg-Fraxe, Hella  
TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF  
FILE REFERENCE: 1487/3  
CURRENT APPLICATION NUMBER: US/10/433,640  
CURRENT FILING DATE: 2003-10-02  
PRIOR APPLICATION NUMBER: PCT/EP01/14610  
PRIOR FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: DE 10061872.3  
PRIOR FILING DATE: 2000-12-12  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 13  
LENGTH: 225  
TYPE: PRT  
ORGANISM: *Diacoboma* sp.  
US-10-433-640-13

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVMREGTVNGHGFIEGSGRPRYGNTTVKLTGKGPLPPAMD 60  
1 MSSKNVKEFMRFKVMREGTVNGHGFIEGSGRPRYGNTTVKLTGKGPLPPAMD 60  
DB 1 MSSKNVKEFMRFKVMREGTVNGHGFIEGSGRPRYGNTTVKLTGKGPLPPAMD 60  
61 LSPQFGSKVYVKKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120

DB 61 LSPQFGSKVYVKKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
DB 121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
QY 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
DB 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 13  
US-10-311-030-7  
Sequence 7, Application US/10311030  
Publication No. US20040171107A1  
GENERAL INFORMATION:  
APPLICANT: Nelson, David  
APPLICANT: Zamiatra, Elize  
APPLICANT: Telen, Roger  
TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS  
FILE REFERENCE: 15916-0320S1  
CURRENT APPLICATION NUMBER: US/10/311,030  
CURRENT FILING DATE: 2002-12-10  
PRIOR APPLICATION NUMBER: PCT/US01/04625  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: US 60/184,732  
PRIOR FILING DATE: 2000-02-23  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic construct  
US-10-311-030-7

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVMREGTVNGHGFIEGSGRPRYGNTTVKLTGKGPLPPAMD 60  
1 MSSKNVKEFMRFKVMREGTVNGHGFIEGSGRPRYGNTTVKLTGKGPLPPAMD 60  
DB 1 MSSKNVKEFMRFKVMREGTVNGHGFIEGSGRPRYGNTTVKLTGKGPLPPAMD 60  
61 LSPQFGSKVYVKKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120  
QY 61 LSPQFGSKVYVKKHPADIPDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFTY 120  
121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
DB 121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
121 KYKFIGVNPSPDGPVWQKKTGWEASTERLYPRDGYLKGELIKALKDKGGHYLVEFKSI 180  
QY 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
DB 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 14  
US-10-845-484-3  
Sequence 3, Application US/10845484  
Publication No. US20040248180A1  
GENERAL INFORMATION:  
APPLICANT: Bullina, Maria  
APPLICANT: Chudakov, Dmitry  
APPLICANT: Lukyanov, Konstantin  
TITLE OF INVENTION: Mutant Chromophores/Fluorophores and  
TITLE OF INVENTION: Methods for Making and Using the Same  
FILE REFERENCE: CLON 092  
CURRENT APPLICATION NUMBER: US/10/845,484  
CURRENT FILING DATE: 2004-05-12  
PRIOR APPLICATION NUMBER: 60/343128  
PRIOR FILING DATE: 2001-12-26

PRIOR APPLICATION NUMBER: PCT/US02/41418  
PRIOR FILING DATE: 2002-12-23  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
US-10-845-484-3

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVTKERFRKVRMEGTNGHEFEIEGEGRPYEGHNTVTKLTGKGGLPFPAMD 60  
DB 1 MRSSKNVTKERFRKVRMEGTNGHEFEIEGEGRPYEGHNTVTKLTGKGGLPFPAMD 60  
QY 61 LSPQOYGSKYVVKRPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQOYGSKYVVKRPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120  
QY 121 KVFEGVNFPSDGPVMOGKKTGMWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
DB 121 KVFEGVNFPSDGPVMOGKKTGMWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
QY 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIYQYERTGRRHHLFL 225  
DB 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIYQYERTGRRHHLFL 225

RESULT 15  
US-10-885-988-12  
Sequence 12, Application US/10885988  
Publication No. US20040259165A1  
GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSIEN, Roger  
TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
FILE REFERENCE: REGEN1530-2  
CURRENT APPLICATION NUMBER: US/10/885,988  
CURRENT FILING DATE: 2004-07-06  
PRIOR APPLICATION NUMBER: US/09/866,538  
PRIOR FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 12  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
US-10-885-988-12

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVTKERFRKVRMEGTNGHEFEIEGEGRPYEGHNTVTKLTGKGGLPFPAMD 60  
DB 1 MRSSKNVTKERFRKVRMEGTNGHEFEIEGEGRPYEGHNTVTKLTGKGGLPFPAMD 60  
QY 61 LSPQOYGSKYVVKRPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQOYGSKYVVKRPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120  
QY 121 KVFEGVNFPSDGPVMOGKKTGMWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
DB 121 KVFEGVNFPSDGPVMOGKKTGMWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
QY 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIYQYERTGRRHHLFL 225  
DB 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIYQYERTGRRHHLFL 225

RESULT 16  
US-10-656-029-22  
Sequence 22, Application US/10656029  
Publication No. US2005000367A1  
GENERAL INFORMATION:  
APPLICANT: VERTEX PHARMACEUTICALS INC.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF  
TITLE OF INVENTION: SCREENING ASSAYS  
FILE REFERENCE: VFI/02-143WO2  
CURRENT APPLICATION NUMBER: US/10/656,029  
CURRENT FILING DATE: 2003-09-05  
PRIOR APPLICATION NUMBER: 60/408,297  
PRIOR FILING DATE: 2002-09-05  
NUMBER OF SEQ ID NOS: 86  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 22  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
US-10-656-029-22

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVTKERFRKVRMEGTNGHEFEIEGEGRPYEGHNTVTKLTGKGGLPFPAMD 60  
DB 1 MRSSKNVTKERFRKVRMEGTNGHEFEIEGEGRPYEGHNTVTKLTGKGGLPFPAMD 60  
QY 61 LSPQOYGSKYVVKRPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQOYGSKYVVKRPADIPDYKLSFPEGFKMERVNNFEDGVTVTQDSSLQDGCFTY 120  
QY 121 KVFEGVNFPSDGPVMOGKKTGMWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
DB 121 KVFEGVNFPSDGPVMOGKKTGMWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
QY 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIYQYERTGRRHHLFL 225  
DB 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTVIYQYERTGRRHHLFL 225

RESULT 17  
US-10-857-622-12  
Sequence 12, Application US/10857622  
Publication No. US20050026234A1  
GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: VIOLIN, Jonathan  
APPLICANT: NEWTON, Alexandra  
APPLICANT: TSIEN, Roger  
TITLE OF INVENTION: Emission Ratiometric Indicators of Phosphorylation By C-Kinase  
FILE REFERENCE: 39754-0891 CPCICP2  
CURRENT APPLICATION NUMBER: US/10/857,622  
CURRENT FILING DATE: 2004-05-28  
PRIOR APPLICATION NUMBER: US 09/865,291  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: US 09/396,003  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: US 08/792,553  
PRIOR FILING DATE: 1997-01-31  
PRIOR APPLICATION NUMBER: US 594,575  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
US-10-857-622-12

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEPMRPRVRNEGTVNGHEFEIEBEGGRPYEGHNTVTKLKTGKGPFPFAMD 60  
DB 1 MRSSKNVKEPMRPRVRNEGTVNGHEFEIEBEGGRPYEGHNTVTKLKTGKGPFPFAMD 60  
QY 61 LSPQFQYSKYVVKHAPADIPDYKLSFPEGFKMERVNFEDGVTVTTQDSSLQDGCFTY 120  
DB 61 LSPQFQYSKYVVKHAPADIPDYKLSFPEGFKMERVNFEDGVTVTTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGPVQWQKTMGWEASTERLYPRDGLVKGELIKALKXDGCHYLVFEXSI 180  
DB 121 KYKFIGVNPSPDGPVQWQKTMGWEASTERLYPRDGLVKGELIKALKXDGCHYLVFEXSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225

## RESULT 18

US-10-505-486-26  
Sequence 26, Application US/10505486  
Publication No. US20050118639A1  
GENERAL INFORMATION:  
APPLICANT: Takeda Chemical Industries, Ltd.  
TITLE OF INVENTION: Determination of a ligand  
FILE REFERENCE: P03-0006PCT  
CURRENT FILING DATE: 2004-08-20  
PRIOR APPLICATION NUMBER: JP 2002-45728  
PRIOR FILING DATE: 2002-02-22  
PRIOR APPLICATION NUMBER: JP 2002-213949  
PRIOR FILING DATE: 2002-07-23  
PRIOR APPLICATION NUMBER: JP 2002-298237  
PRIOR FILING DATE: 2002-10-11  
NUMBER OF SEQ ID NOS: 233  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
US-10-505-486-26

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEPMRPRVRNEGTVNGHEFEIEBEGGRPYEGHNTVTKLKTGKGPFPFAMD 60  
DB 1 MRSSKNVKEPMRPRVRNEGTVNGHEFEIEBEGGRPYEGHNTVTKLKTGKGPFPFAMD 60  
QY 61 LSPQFQYSKYVVKHAPADIPDYKLSFPEGFKMERVNFEDGVTVTTQDSSLQDGCFTY 120  
DB 61 LSPQFQYSKYVVKHAPADIPDYKLSFPEGFKMERVNFEDGVTVTTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGPVQWQKTMGWEASTERLYPRDGLVKGELIKALKXDGCHYLVFEXSI 180  
DB 121 KYKFIGVNPSPDGPVQWQKTMGWEASTERLYPRDGLVKGELIKALKXDGCHYLVFEXSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225

## RESULT 19

US-10-844-064A-2  
Sequence 2, Application US/10844064A  
Publication No. US20050149994A1  
GENERAL INFORMATION:  
APPLICANT: Bevis, Brooke  
APPLICANT: Glick, Benjamin  
TITLE OF INVENTION: RAPIDLY MATURING FLUORESCENT PROTEINS AND METHODS FOR USING THE S

FILE REFERENCE: 092234-9006  
CURRENT APPLICATION NUMBER: US/10/844, 064A  
CURRENT FILING DATE: 2004-05-11  
PRIOR APPLICATION NUMBER: PCT/US02/40539  
PRIOR FILING DATE: 2002-12-18  
PRIOR APPLICATION NUMBER: US 60/341,723  
PRIOR FILING DATE: 2001-12-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma  
US-10-844-064A-2

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEPMRPRVRNEGTVNGHEFEIEBEGGRPYEGHNTVTKLKTGKGPFPFAMD 60  
DB 1 MRSSKNVKEPMRPRVRNEGTVNGHEFEIEBEGGRPYEGHNTVTKLKTGKGPFPFAMD 60  
QY 61 LSPQFQYSKYVVKHAPADIPDYKLSFPEGFKMERVNFEDGVTVTTQDSSLQDGCFTY 120  
DB 61 LSPQFQYSKYVVKHAPADIPDYKLSFPEGFKMERVNFEDGVTVTTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNPSPDGPVQWQKTMGWEASTERLYPRDGLVKGELIKALKXDGCHYLVFEXSI 180  
DB 121 KYKFIGVNPSPDGPVQWQKTMGWEASTERLYPRDGLVKGELIKALKXDGCHYLVFEXSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIYEQYERTGRHHLFL 225

## RESULT 20

US-10-931-304-1  
Sequence 1, Application US/10931304  
Publication No. US20050196768A1  
GENERAL INFORMATION:  
APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
FILE REFERENCE: 39754-0831CP2CP3  
CURRENT APPLICATION NUMBER: US/10/931,304  
CURRENT FILING DATE: 2004-08-30  
PRIOR APPLICATION NUMBER: 10/209,208  
PRIOR FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Discosoma sp.  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(225)  
OTHER INFORMATION: wild-type DsRed  
US-10-931-304-1

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





Db 342 LSPQFQYSGKYVVKHPADI PDYKCLSPFEGFKMERVNFEDGGVTVTVTQDSSLQDGCFTY 401  
 QY 121 KYKFGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
 Db 402 KYKFGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 461  
 QY 181 YMAKRPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 Db 462 YMAKRPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 506

RESULT 24  
 US-10-343-977-2  
 ; Sequence 2, Application US/10343977  
 ; Publication No. US20050112682A9  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kuhlmann, Rene  
 ; APPLICANT: Koltermann, Andre  
 ; APPLICANT: Ketting, Ulrich  
 ; APPLICANT: Schwille, Petra  
 ; TITLE OF INVENTION: Dual coloured fluorimetric protease assay  
 ; FILE REFERENCE: 023425us/0H/ml  
 ; CURRENT APPLICATION NUMBER: US/10/343,977  
 ; CURRENT FILING DATE: 2003-02-06  
 ; PRIOR APPLICATION NUMBER: PCT/EP01/09112  
 ; PRIOR FILING DATE: 2001-08-07  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 506  
 ; TYPE: PRT  
 ; ORGANISM: Artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Autofluorescent fusion protein derived from GFP variants of  
 ; OTHER INFORMATION: e.g. Aequorea victoria and/or Discoma sp.  
 US-10-343-977-2

Query Match 100.0%; Score 1214; DB 5; Length 506;  
 Best Local Similarity 100.0%; Pred. No. 2e-113;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSSSKVNIKEFMRFKVMEGTGNGHFEIIEGEGRPYEGHNTVTKLVTKGGLPFPAMD 60  
 Db 282 MSSSKVNIKEFMRFKVMEGTGNGHFEIIEGEGRPYEGHNTVTKLVTKGGLPFPAMD 341  
 QY 61 LSPQFQYSGKYVVKHPADI PDYKCLSPFEGFKMERVNFEDGGVTVTVTQDSSLQDGCFTY 120  
 Db 342 LSPQFQYSGKYVVKHPADI PDYKCLSPFEGFKMERVNFEDGGVTVTVTQDSSLQDGCFTY 401  
 QY 121 KYKFGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
 Db 402 KYKFGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 461  
 QY 181 YMAKRPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 Db 462 YMAKRPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 506

RESULT 25  
 US-10-343-977-3  
 ; Sequence 3, Application US/10343977  
 ; Publication No. US20040014134A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kuhlmann, Rene  
 ; APPLICANT: Koltermann, Andre  
 ; APPLICANT: Ketting, Ulrich  
 ; APPLICANT: Schwille, Petra  
 ; TITLE OF INVENTION: Dual coloured fluorimetric protease assay  
 ; FILE REFERENCE: 023425us/0H/ml  
 ; CURRENT APPLICATION NUMBER: US/10/343,977  
 ; CURRENT FILING DATE: 2003-02-06  
 ; PRIOR APPLICATION NUMBER: PCT/EP01/09112

; PRIOR FILING DATE: 2001-08-07  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 547  
 ; TYPE: PRT  
 ; ORGANISM: Artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Autofluorescent fusion protein derived from GFP variants of  
 ; OTHER INFORMATION: e.g. Aequorea victoria and/or Discoma sp.  
 US-10-343-977-3

Query Match 100.0%; Score 1214; DB 4; Length 547;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-113;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSSSKVNIKEFMRFKVMEGTGNGHFEIIEGEGRPYEGHNTVTKLVTKGGLPFPAMD 60  
 Db 323 MSSSKVNIKEFMRFKVMEGTGNGHFEIIEGEGRPYEGHNTVTKLVTKGGLPFPAMD 382  
 QY 61 LSPQFQYSGKYVVKHPADI PDYKCLSPFEGFKMERVNFEDGGVTVTVTQDSSLQDGCFTY 120  
 Db 383 LSPQFQYSGKYVVKHPADI PDYKCLSPFEGFKMERVNFEDGGVTVTVTQDSSLQDGCFTY 442  
 QY 121 KYKFGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
 Db 443 KYKFGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 502  
 QY 181 YMAKRPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
 Db 503 YMAKRPVOLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 547

RESULT 26  
 US-10-343-977-3  
 ; Sequence 3, Application US/10343977  
 ; Publication No. US20050112682A9  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kuhlmann, Rene  
 ; APPLICANT: Koltermann, Andre  
 ; APPLICANT: Ketting, Ulrich  
 ; APPLICANT: Schwille, Petra  
 ; TITLE OF INVENTION: Dual coloured fluorimetric protease assay  
 ; FILE REFERENCE: 023425us/0H/ml  
 ; CURRENT APPLICATION NUMBER: US/10/343,977  
 ; CURRENT FILING DATE: 2003-02-06  
 ; PRIOR APPLICATION NUMBER: PCT/EP01/09112  
 ; PRIOR FILING DATE: 2001-08-07  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 547  
 ; TYPE: PRT  
 ; ORGANISM: Artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Autofluorescent fusion protein derived from GFP variants of  
 ; OTHER INFORMATION: e.g. Aequorea victoria and/or Discoma sp.  
 US-10-343-977-3

Query Match 100.0%; Score 1214; DB 5; Length 547;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-113;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSSSKVNIKEFMRFKVMEGTGNGHFEIIEGEGRPYEGHNTVTKLVTKGGLPFPAMD 60  
 Db 323 MSSSKVNIKEFMRFKVMEGTGNGHFEIIEGEGRPYEGHNTVTKLVTKGGLPFPAMD 382  
 QY 61 LSPQFQYSGKYVVKHPADI PDYKCLSPFEGFKMERVNFEDGGVTVTVTQDSSLQDGCFTY 120  
 Db 383 LSPQFQYSGKYVVKHPADI PDYKCLSPFEGFKMERVNFEDGGVTVTVTQDSSLQDGCFTY 442  
 QY 121 KYKFGVNPSPDGPVWQKKTGMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180





```
/ CURRENT FILING DATE: 2002-12-09
/ PRIOR APPLICATION NUMBER: 60/211,607
/ PRIOR FILING DATE: 2000-06-14
/ PRIOR APPLICATION NUMBER: PCT/US01/19097
/ PRIOR FILING DATE: 2001-06-13
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-4

Query Match          99.4%; Score 1207; DB 4; Length 225;
Best Local Similarity 99.1%; Pred. No. 3.3e-113;
Matches 223; Conservative 1; Mismatches 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHFEIIEGEGRPRYEGHNTVTLKVTKGGLPFPANDI 60
    |||
DB 1 MRSSKNVIKEFMRFKVRMEGTVNGHFEIIEGEGRPRYEGHNTVTLKVTKGGLPFPANDI 60

QY 61 LSPQFQYGSKYVYVGHADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120
    |||
DB 61 LSPQFQYGSKYVYVGHADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120

QY 121 KVFIFGVNFPDSGVVMQKTMGWEASTERYLPDGVYLKGEIHKALKDKGGHYLVFPKSI 180
    |||
DB 121 KVFIFGVNFPDSGVVMQKTMGWEASTERYLPDGVYLKGEIHKALKDKGGHYLVFPKSI 180

QY 181 YMAKKPVQLPGYIVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225
    |||
DB 181 YMAKKPVQLPGYIVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 34
US-10-931-304-20
/ Sequence 20, Application US/10931304
/ Publication No. US20050196768A1
/ GENERAL INFORMATION:
/ APPLICANT: Telen, Roger
/ APPLICANT: Campbell, Robert
/ APPLICANT: Baird, Geoffrey
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: 39754-0831CP2CP3
/ CURRENT APPLICATION NUMBER: US/10/931,304
/ CURRENT FILING DATE: 2004-08-30
/ PRIOR APPLICATION NUMBER: 10/209,208
/ PRIOR FILING DATE: 2002-07-29
/ PRIOR APPLICATION NUMBER: 10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 20
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Dared with I125R
US-10-931-304-20

Query Match          99.4%; Score 1207; DB 5; Length 225;
Best Local Similarity 99.6%; Pred. No. 3.3e-113;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHFEIIEGEGRPRYEGHNTVTLKVTKGGLPFPANDI 60
    |||
```

```
DB 1 MRSSKNVIKEFMRFKVRMEGTVNGHFEIIEGEGRPRYEGHNTVTLKVTKGGLPFPANDI 60

QY 61 LSPQFQYGSKYVYVGHADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120
    |||
DB 61 LSPQFQYGSKYVYVGHADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120

QY 121 KVFIFGVNFPDSGVVMQKTMGWEASTERYLPDGVYLKGEIHKALKDKGGHYLVFPKSI 180
    |||
DB 121 KVFIFGVNFPDSGVVMQKTMGWEASTERYLPDGVYLKGEIHKALKDKGGHYLVFPKSI 180

QY 181 YMAKKPVQLPGYIVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225
    |||
DB 181 YMAKKPVQLPGYIVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 35
US-10-314-936-2
/ Sequence 2, Application US/10314936
/ Publication No. US20040110225A1
/ GENERAL INFORMATION:
/ APPLICANT: Gibbs, Patrick D.L.
/ APPLICANT: Carter, Robert W.
/ APPLICANT: Schmale, Michael C.
/ TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES
/ FILE REFERENCE: 638,004
/ CURRENT APPLICATION NUMBER: US/10/314,936
/ CURRENT FILING DATE: 2002-12-09
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 236
/ TYPE: PRT
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: mutant red fluorescent protein
US-10-314-936-2

Query Match          98.5%; Score 1196; DB 4; Length 236;
Best Local Similarity 98.2%; Pred. No. 4.6e-112;
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHFEIIEGEGRPRYEGHNTVTLKVTKGGLPFPANDI 60
    |||
DB 1 MRSSKNVIKEFMRFKVRMEGTVNGHFEIIEGEGRPRYEGHNTVTLKVTKGGLPFPANDI 60

QY 61 LSPQFQYGSKYVYVGHADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120
    |||
DB 61 LSPQFQYGSKYVYVGHADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120

QY 121 KVFIFGVNFPDSGVVMQKTMGWEASTERYLPDGVYLKGEIHKALKDKGGHYLVFPKSI 180
    |||
DB 121 KVFIFGVNFPDSGVVMQKTMGWEASTERYLPDGVYLKGEIHKALKDKGGHYLVFPKSI 180

QY 181 YMAKKPVQLPGYIVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225
    |||
DB 181 YMAKKPVQLPGYIVYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 36
US-10-314-936-4
/ Sequence 4, Application US/10314936
/ Publication No. US20040110225A1
/ GENERAL INFORMATION:
/ APPLICANT: Gibbs, Patrick D.L.
/ APPLICANT: Carter, Robert W.
/ APPLICANT: Schmale, Michael C.
/ TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES
/ FILE REFERENCE: 638,004
/ CURRENT APPLICATION NUMBER: US/10/314,936
/ CURRENT FILING DATE: 2002-12-09
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 4
```

LENGTH: 236  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 OTHER INFORMATION: mutant red fluorescent protein  
 US-10-314-936-4

Query Match 98.5%; Score 1196; DB 6; Length 236;  
 Best Local Similarity 98.2%; Pred. No. 4,6e-112;  
 Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVRMEGTNGHEFEIEGEGRPYEGHNTVLKTKTGGLPFPAMD 60  
 DB 1 MSCSKVNIKEFMRFKVRMEGTNGHEFEIEGEGRPYEGHNTVLKTKTGGLPFPAMD 60  
 QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGGVTVTODSLQDGCFTY 120  
 DB 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGGVTVTODSLQDGCFTY 120  
 QY 121 KYKFGVNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
 DB 121 KYKFGVNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKTI 180  
 QY 181 YMAKPPVQLPGYYVYDSKLDITSNEDYTIIVEQYERTGRHHLFL 225  
 DB 181 YMAKPPVQLPGYYVYDSKLDITSNEDYTIIVEQYERTGRHHLFL 225

## RESULT 37

US-11-021-014-2,  
 Sequence 2, Application US/11021014  
 Publication No. US20050100954A1  
 GENERAL INFORMATION:

APPLICANT: Gibbs, Patrick D.L.  
 APPLICANT: Carter, Robert W.  
 APPLICANT: Schmale, Michael C.  
 TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
 FILE REFERENCE: 638.004  
 CURRENT APPLICATION NUMBER: US/11/021,014  
 CURRENT FILING DATE: 2004-12-23  
 PRIOR APPLICATION NUMBER: US/10/314,936  
 PRIOR FILING DATE: 2002-12-09  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 2  
 LENGTH: 236  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 OTHER INFORMATION: mutant red fluorescent protein  
 US-11-021-014-2

Query Match 98.5%; Score 1196; DB 6; Length 236;  
 Best Local Similarity 98.2%; Pred. No. 4,6e-112;  
 Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVRMEGTNGHEFEIEGEGRPYEGHNTVLKTKTGGLPFPAMD 60  
 DB 1 MSCSKVNIKEFMRFKVRMEGTNGHEFEIEGEGRPYEGHNTVLKTKTGGLPFPAMD 60  
 QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGGVTVTODSLQDGCFTY 120  
 DB 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGGVTVTODSLQDGCFTY 120  
 QY 121 KYKFGVNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
 DB 121 KYKFGVNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKTI 180  
 QY 181 YMAKPPVQLPGYYVYDSKLDITSNEDYTIIVEQYERTGRHHLFL 225  
 DB 181 YMAKPPVQLPGYYVYDSKLDITSNEDYTIIVEQYERTGRHHLFL 225

## RESULT 38

US-10-724-178-12  
 Sequence 4, Application US/11021014  
 Publication No. US20050100954A1  
 GENERAL INFORMATION:

APPLICANT: Gibbs, Patrick D.L.  
 APPLICANT: Carter, Robert W.  
 APPLICANT: Schmale, Michael C.  
 TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
 FILE REFERENCE: 638.004  
 CURRENT APPLICATION NUMBER: US/11/021,014  
 CURRENT FILING DATE: 2004-12-23  
 PRIOR APPLICATION NUMBER: US/10/314,936  
 PRIOR FILING DATE: 2002-12-09  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 4  
 LENGTH: 236  
 TYPE: PRT  
 ORGANISM: Artificial  
 FEATURE:  
 OTHER INFORMATION: mutant red fluorescent protein  
 US-11-021-014-4

Query Match 98.5%; Score 1196; DB 6; Length 236;  
 Best Local Similarity 98.2%; Pred. No. 4,6e-112;  
 Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMRFKVRMEGTNGHEFEIEGEGRPYEGHNTVLKTKTGGLPFPAMD 60  
 DB 1 MSCSKVNIKEFMRFKVRMEGTNGHEFEIEGEGRPYEGHNTVLKTKTGGLPFPAMD 60  
 QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGGVTVTODSLQDGCFTY 120  
 DB 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGGVTVTODSLQDGCFTY 120  
 QY 121 KYKFGVNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSI 180  
 DB 121 KYKFGVNPSPDGPVWQKKTWMEASTERLYPRDGLKGEIHKALKDGGHYLVEFKTI 180  
 QY 181 YMAKPPVQLPGYYVYDSKLDITSNEDYTIIVEQYERTGRHHLFL 225  
 DB 181 YMAKPPVQLPGYYVYDSKLDITSNEDYTIIVEQYERTGRHHLFL 225

## RESULT 39

US-10-724-178-12  
 Sequence 12, Application US/10724178  
 Publication No. US20040137528A1  
 GENERAL INFORMATION:

APPLICANT: Odysey Thera, Inc.  
 APPLICANT: Michnick, Stephen  
 APPLICANT: Macdonald, Marilee  
 APPLICANT: Lamerdin, Jane  
 TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT  
 FILE REFERENCE: ODDY007  
 CURRENT APPLICATION NUMBER: US/10/724,178  
 CURRENT FILING DATE: 2003-12-01  
 PRIOR APPLICATION NUMBER: US 60/461,133  
 PRIOR FILING DATE: 2003-04-09  
 NUMBER OF SEQ ID NOS: 1067  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO 12  
 LENGTH: 226  
 TYPE: PRT  
 ORGANISM: Anthozoa discosoma;  
 NAME/KEY: misc\_feature  
 OTHER INFORMATION: Discosoma red ("dr") FP583  
 PUBLICATION INFORMATION:  
 AUTHORS: Matz, M.V. et al  
 TITLE: Fluorescent proteins from nonbioluminescent Anthozoa species

JOURNAL: Nat. Biotech.  
; VOLUME: 17  
; ISSUE: 10  
; PAGES: 969-73  
; DATE: 1999-  
; RELEVANT RESIDUES: (1)..(226)  
US-10-724-178-12

Query Match 98.5%; Score 1195.5; DB 4; Length 226;  
Best Local Similarity 99.1%; Pred. No. 4.9e-112;  
Matches 224; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MRSSKNVKEFRPFYKRMGTNGHGFIEEGEGRPYEGHNTVTLKTKGGPLPFAMD 60  
DB 1 MRSSKNVKEFRPFYKRMGTNGHGFIEEGEGRPYEGHNTVTLKTKGGPLPFAMD 60  
QY 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGVVTVTQDSSLQDGCFTY 120  
DB 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGVVTVTQDSSLQDGCFTY 120  
QY 121 KVFRTGVNPPSPGPMOKKTWMEASTERYLPDGVLTGKETHKALKDGGHYLVEFKSI 180  
DB 121 KVFRTGVNPPSPGPMOKKTWMEASTERYLPDGVLTGKETHKALKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRRHFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRRHFL 226

RESULT 40  
US-10-315-920-6  
; Sequence 6, Application US/10315920  
; Publication No. US20030175809A1  
; GENERAL INFORMATION:  
; APPLICANT: Fradkov, Arcady Fedorovich  
; APPLICANT: Terelikh, Alexey  
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS  
; TITLE OF INVENTION: FOR THEIR USE  
; FILE REFERENCE: CLON-077C1P  
; CURRENT APPLICATION NUMBER: US/10/315,920  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: 60/211,607  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: PCT/US01/19097  
; PRIOR FILING DATE: 2001-06-13  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: variant of sequence from *Dicosoma* sp.  
US-10-315-920-6

Query Match 98.1%; Score 1191; DB 4; Length 225;  
Best Local Similarity 97.8%; Pred. No. 1.4e-111;  
Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRPFYKRMGTNGHGFIEEGEGRPYEGHNTVTLKTKGGPLPFAMD 60  
DB 1 MASSENVITFERPFYKRMGTNGHGFIEEGEGRPYEGHNTVTLKTKGGPLPFAMD 60  
QY 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGVVTVTQDSSLQDGCFTY 120  
DB 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGVVTVTQDSSLQDGCFTY 120  
QY 121 KVFRTGVNPPSPGPMOKKTWMEASTERYLPDGVLTGKETHKALKDGGHYLVEFKSI 180  
DB 121 KVFRTGVNPPSPGPMOKKTWMEASTERYLPDGVLTGKETHKALKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRRHFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRRHFL 225

DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRRHFL 225

RESULT 41  
US-10-442-148A-7  
; Sequence 7, Application US/10442148A  
; Publication No. US20040014242A1  
; GENERAL INFORMATION:  
; APPLICANT: IWAKURA, MASAHIRO  
; APPLICANT: HIROTA, KIYONORI  
; TITLE OF INVENTION: PROCESS FOR IMMOBILIZING ORIENTATION-CONTROLLED PROTEIN AND  
; TITLE OF INVENTION: PROCESS FOR ARRAYING AND IMMOBILIZING PROTEIN USING THE SAME  
; FILE REFERENCE: 04583.0103-00000  
; CURRENT APPLICATION NUMBER: US/10/442,148A  
; CURRENT FILING DATE: 2003-05-21  
; PRIOR APPLICATION NUMBER: JP 2002-148950  
; PRIOR FILING DATE: 2002-05-23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein sequence  
US-10-442-148A-7

Query Match 97.7%; Score 1186; DB 4; Length 225;  
Best Local Similarity 97.3%; Pred. No. 4.4e-111;  
Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRPFYKRMGTNGHGFIEEGEGRPYEGHNTVTLKTKGGPLPFAMD 60  
DB 1 MASSENVITFERPFYKRMGTNGHGFIEEGEGRPYEGHNTVTLKTKGGPLPFAMD 60  
QY 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGVVTVTQDSSLQDGCFTY 120  
DB 61 LSPFOYGSKYVVKHPADIPDYKLSFPEGFKMERVNFEDGVVTVTQDSSLQDGCFTY 120  
QY 121 KVFRTGVNPPSPGPMOKKTWMEASTERYLPDGVLTGKETHKALKDGGHYLVEFKSI 180  
DB 121 KVFRTGVNPPSPGPMOKKTWMEASTERYLPDGVLTGKETHKALKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRRHFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRRHFL 225

RESULT 42  
US-10-442-148A-8  
; Sequence 8, Application US/10442148A  
; Publication No. US20040014242A1  
; GENERAL INFORMATION:  
; APPLICANT: IWAKURA, MASAHIRO  
; APPLICANT: HIROTA, KIYONORI  
; TITLE OF INVENTION: PROCESS FOR IMMOBILIZING ORIENTATION-CONTROLLED PROTEIN AND  
; TITLE OF INVENTION: PROCESS FOR ARRAYING AND IMMOBILIZING PROTEIN USING THE SAME  
; FILE REFERENCE: 04583.0103-00000  
; CURRENT APPLICATION NUMBER: US/10/442,148A  
; CURRENT FILING DATE: 2003-05-21  
; PRIOR APPLICATION NUMBER: JP 2002-148950  
; PRIOR FILING DATE: 2002-05-23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 8  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein sequence  
US-10-442-148A-8

Query Match 97.7%; Score 1186; DB 4; Length 239;

Best Local Similarity 97.3%; Pred. No. 4.8e-111;  
Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy	1	MASSINQVLKEPRFRVYRMWGVNNGHEPFI	1	BEESGSGRPYBEGNNYTKLKVTGCGSL	PF	AMDI	60
	1	MASSEVLTTEFRFRVYRMWGVNNGHEPFI	1	BEESGSGRPYBEGNNYTKLKVTGCGSL	PF	AMDI	60
Db							
Qy	61	LSPOEQYSGSKVYVKKHAPADIPDYKLS	61	SPBEGGKMERVNNFBDGCVYVITODSSLDGCGFIY			120
Db	61	LSPOEQYSGSKVYVKKHAPADIPDYKLS	61	SPBEGGKMERVNNFBDGCVYVITODSSLDGCGFIY			120
Qy	121	KXKFTGVNPPSDGPRWOKKTIMGWEASTERL	121	YPRDGVTLKGEIHKALKKODGHTLVVERKSI			180
Db	121	KXKFTGVNPPSDGPRWOKKTIMGWEASTERL	121	YPRDGVTLKGEIHKALKKODGHTLVVERKSI			180
Qy	181	YMAKKRPVOLPGYVYVDSKLDITSHNEHDYI	181	VEQYVETRGHTHLP			225
Db	181	YMAKKRPVOLPGYVYVDSKLDITSHNEHDYI	181	VEQYVETRGHTHLP			225

RESULT 43

```

US-10-121-258-4
/ Sequence 4, Application US/10121258
/ Publication No. US20030059835A1
/ GENERAL INFORMATION:
/ APPLICANT: Tsien, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: UC083.1CP2CB1
/ CURRENT APPLICATION NUMBER: US/10/121,258
/ CURRENT FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Dered polypeptide variant "I1"
/ US-10-121-258-4

```

Query match	95.6%	Score 1160	DB 4	Length 225
-------------	-------	------------	------	------------

[illegible]

RESULT 44  
US-10-931-304-4  
; Sequence 4, Application US/109313304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger

```

? APPLICANT: Campbell, Robert
? APPLICANT: Baird, Geoffrey
? TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
? TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
? FILE REFERENCE: 39754-0831CP2CP3
? CURRENT APPLICATION NUMBER: US/10/931.304
? PRIOR FILING DATE: 2004-08-30
? PRIOR APPLICATION NUMBER: 10/209,208
? PRIOR FILING DATE: 2002-07-29
? PRIOR APPLICATION NUMBER: 10/121,258
? PRIOR FILING DATE: 2002-04-10
? PRIOR APPLICATION NUMBER: 09/866,538
? PRIOR FILING DATE: 2001-05-24
? PRIOR APPLICATION NUMBER: 09/794,308
? PRIOR FILING DATE: 2001-02-26
? NUMBER OF SEQ ID NOS: 110
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 4
? LENGTH: 225
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Dared polypeptide variant "T1"
? US-10-931-304-4

```

Query Match	Score	DB 5;	Length
95.6%	1160;	DB 5;	2257

Qy	1	MSSKXVITEPFRFLYVRBGTATNGHEPFI	BEBSGGR	PEGHNTYTKL	LYTKGSP	LPFAMD	60		
		1							
		1	MSSEVIEPEFRFKRMEGVSNGHEPFI	BEBSGGR	PEEGTQTKL	LYTKGSP	LPFAMD	60	
Db		1	MSSEVIEPEFRFKRMEGVSNGHEPFI	BEBSGGR	PEEGTQTKL	LYTKGSP	LPFAMD	60	
Qy		61	LSQFOYSGSKVYVYKHPADI	PDYKKL	SPFEGFKMERVNF	DGVVTVTODS	LDGCFIY	120	
		61	LSQFOYSGSKVYVYKHPADI	PDYKKL	SPFEGFKMERVNF	DGVVTVTODS	LDGCFIY	120	
Db		61	LSQFOYSGSKVYVYKHPADI	PDYKKL	SPFEGFKMERVNF	DGVVTVTODS	LDGCFIY	120	
Qy		121	KVFIGNVFPSSDGPVWOKKTWGMBA	STERL	YPRDGYL	KGELHKL	KLKDGGHYLV	ERSI	180
		121	KVFIGNVFPSSDGPVWOKKTWGMBA	STERL	YPRDGYL	KGELHKL	KLKDGGHYLV	ERSI	180
Db		121	KVFIGNVFPSSDGPVWOKKTWGMBA	STERL	YPRDGYL	KGELHKL	KLKDGGHYLV	ERSI	180
Qy		181	YMAKKEPVOLPGYVYVDSKLDITS	SHNBDITYE	QVYERTEGR	HNHFL	FL	225	
		181	YMAKKEPVOLPGYVYVDSKLDITS	SHNBDITYE	QVYERTEGR	HNHFL	FL	225	
Db		181	YMAKKEPVOLPGYVYVDSKLDITS	SHNBDITYE	QVYERTEGR	HNHFL	FL	225	

RESULT 45  
10-121

```

US-10-121-258-24
; Sequence 24, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Cambell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1c2cP1
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DsRed polypeptide variant "r1" with 1125R mutation
US-10-121-258-24
Query Match          95.0%; Score 1153; DB 4; Length 225;

```





```

Db      4 SSEDVYKEFRMRKRVMEGVNHEFEIEBEGEGRPVEGHTQAKLKYTKGGPLPFAMDIIS 63
Qy      63 POFQYGSKYVYVHPADIPYKLSFPEGFKMERVMNFEDGGVVTYQDSSLQDGCPIYKY 122
Db      64 POFQYGSKAYVYHGPADIPYKLSFPEGFKMERVMNFEDGGVVTYQDSSLQDGTILYKY 123
Qy      123 KPIGVNFPDSGPMQKKTWGMWASTERLYPRDGVLYKGEIHKALKKDGSHYLVFESKIYM 182
Db      124 KRGNTFPDPGPMQKKTWGMWASTERLYPRDGVLYKGEIHKALKKDGSHYLVFESKIYM 183
Qy      183 AKKPVLPGYVYVDSKLDITSHNEDYTIYEYERTEGRHHLFL 225
Db      184 AKKPVLPGYVYVDTKLDITSHNEDYTIYEYERSEGRHHLFL 226

```

## RESULT 49

```

US-10-931-304-6
/ Sequence 6, Application US/10931304
/ Publication No. US20050196768A1
/ GENERAL INFORMATION:

```

```

/ APPLICANT: Telen, Roger
/ APPLICANT: Campbell, Robert
/ APPLICANT: Baild, Geoffrey
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ FILE REFERENCE: 39754-0831CP2CP3
/ CURRENT APPLICATION NUMBER: US/10/931,304
/ CURRENT FILING DATE: 2004-08-30
/ PRIOR APPLICATION NUMBER: 10/209,208
/ PRIOR FILING DATE: 2002-07-29
/ PRIOR APPLICATION NUMBER: 10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6
/ LENGTH: 226
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Dared polypeptide variant "dimer2"
US-10-931-304-6

```

```

Query Match      92.3%; Score 1121; DB 5; Length 226;
Best Local Similarity 92.8%; Pred. No. 1.6e-104;
Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy      3 SSKNVYKEFRMRKRVMEGVNHEFEIEBEGEGRPVEGHTQAKLKYTKGGPLPFAMDIIS 62
Db      4 SSEDVYKEFRMRKRVMEGVNHEFEIEBEGEGRPVEGHTQAKLKYTKGGPLPFAMDIIS 63
Qy      63 POFQYGSKYVYVHPADIPYKLSFPEGFKMERVMNFEDGGVVTYQDSSLQDGCPIYKY 122
Db      64 POFQYGSKAYVYHGPADIPYKLSFPEGFKMERVMNFEDGGVVTYQDSSLQDGTILYKY 123
Qy      123 KPIGVNFPDSGPMQKKTWGMWASTERLYPRDGVLYKGEIHKALKKDGSHYLVFESKIYM 182
Db      124 KRGNTFPDPGPMQKKTWGMWASTERLYPRDGVLYKGEIHKALKKDGSHYLVFESKIYM 183
Qy      183 AKKPVLPGYVYVDSKLDITSHNEDYTIYEYERTEGRHHLFL 225
Db      184 AKKPVLPGYVYVDTKLDITSHNEDYTIYEYERSEGRHHLFL 226

```

## RESULT 50

```

US-11-052-001-6
/ Sequence 6, Application US/11052001
/ Publication No. US2005023356A1
/ GENERAL INFORMATION:

```

```

/ APPLICANT: JONES, JOSHUA T.
/ APPLICANT: HAHN, ANGELA TERESA
/ APPLICANT: MEYER, TOBIAS
/ TITLE OF INVENTION: LIVE-CELL BIOSENSOR POLYPEPTIDES AND
/ FILE REFERENCE: STAN-385
/ CURRENT APPLICATION NUMBER: US/11/052,001
/ CURRENT FILING DATE: 2005-02-03
/ PRIOR APPLICATION NUMBER: 60/541,297
/ PRIOR FILING DATE: 2004-02-03
/ NUMBER OF SEQ ID NOS: 17
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6
/ LENGTH: 602
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: fluorescent fusion protein
US-11-052-001-6

```

```

Query Match      92.3%; Score 1121; DB 6; Length 602;
Best Local Similarity 92.8%; Pred. No. 6.1e-104;
Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy      3 SSKNVYKEFRMRKRVMEGVNHEFEIEBEGEGRPVEGHTQAKLKYTKGGPLPFAMDIIS 62
Db      142 SSEDVYKEFRMRKRVMEGVNHEFEIEBEGEGRPVEGHTQAKLKYTKGGPLPFAMDIIS 201
Qy      63 POFQYGSKYVYVHPADIPYKLSFPEGFKMERVMNFEDGGVVTYQDSSLQDGCPIYKY 122
Db      202 POFQYGSKAYVYHGPADIPYKLSFPEGFKMERVMNFEDGGVVTYQDSSLQDGTILYKY 261
Qy      123 KPIGVNFPDSGPMQKKTWGMWASTERLYPRDGVLYKGEIHKALKKDGSHYLVFESKIYM 182
Db      262 KRGNTFPDPGPMQKKTWGMWASTERLYPRDGVLYKGEIHKALKKDGSHYLVFESKIYM 321
Qy      183 AKKPVLPGYVYVDSKLDITSHNEDYTIYEYERTEGRHHLFL 225
Db      322 AKKPVLPGYVYVDTKLDITSHNEDYTIYEYERSEGRHHLFL 364

```

## RESULT 51

```

US-10-006-922-44
/ Sequence 44, Application US/10006922
/ Publication No. US20020197676A1
/ GENERAL INFORMATION:

```

```

/ APPLICANT: Lukyanov, Sergey A.
/ APPLICANT: Pradkov, Arcady F.
/ APPLICANT: Labae, Yulii A.
/ APPLICANT: Matz, Mikhail V.
/ APPLICANT: Tersikh, Alexey
/ TITLE OF INVENTION: No. US20020197676A1 Chromophores/fluorophores and
/ FILE REFERENCE: C10N-035CIP
/ CURRENT APPLICATION NUMBER: US/10/006,922
/ CURRENT FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 09/120,330
/ PRIOR FILING DATE: 1998-12-11
/ PRIOR APPLICATION NUMBER: 09/457,898
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/458,144
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/458,477
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/457,556
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/444,338
/ PRIOR FILING DATE: 1999-11-19
/ NUMBER OF SEQ ID NOS: 46
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 44
/ LENGTH: 225
/ TYPE: PRT

```

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: hybrid construct  
US-10-006-922-44

Query Match 92.2%; Score 1119; DB 4; Length 225;  
Best Local Similarity 91.6%; Pred. No. 2.5e-104;  
Matches 206; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSSKNVKEKFRFKVMEGTVNGHBEFIEGEGRPYEGHNTVTKYTKGGPLFPADI 60  
DB 1 MSSKNVKEKFRFKVMEGTVNGHBEFIEGEGRPYEGHNSVTLMTYKGGPLFPADI 60  
QY 61 LSPQOYSGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQOYSGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKKTGMWEASTERYPRDGLKGEIHKALKDKGHHYVEFKSI 180  
DB 121 KYKFIGVNFPSDGPVMQKKTGMWEASTERYPRDGLKGDIMALRLBGGHHYVEFKSI 180  
QY 181 YMAKPPVQLPGYVYVDSKLDITSHNEDYTIYQYERTEGRHHLFL 225  
DB 181 YMAKPPVQLPGYVYVDSKLDITSHNEDYTIYQYERTEGRHHLFL 225

## RESULT 52

US-10-081-864-12  
Sequence 12, Application US/10081864  
Publication No. US2003002287A1

## GENERAL INFORMATION:

APPLICANT: Lukyanov, Sergey  
APPLICANT: Lukyanov, Konstantin  
APPLICANT: Yanushevich, Yuriy  
APPLICANT: Savitsky, Alexander  
APPLICANT: Pradkov, Arcady  
TITLE OF INVENTION: No. US2003002287A1 Aggregating Fluorescent Proteins and  
TITLE OF INVENTION: Methods for Using the Same  
FILE REFERENCE: CLON-067  
CURRENT FILING DATE: 2002-06-19  
PRIOR FILING DATE: 2002-06-19  
PRIOR FILING DATE: 2001-12-04  
PRIOR FILING DATE: 2001-12-04  
PRIOR FILING DATE: 2001-02-21  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: hybrid protein  
US-10-081-864-12

Query Match 92.2%; Score 1119; DB 4; Length 225;  
Best Local Similarity 91.6%; Pred. No. 2.5e-104;  
Matches 206; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 1 MSSKNVKEKFRFKVMEGTVNGHBEFIEGEGRPYEGHNTVTKYTKGGPLFPADI 60  
DB 1 MSSKNVKEKFRFKVMEGTVNGHBEFIEGEGRPYEGHNSVTLMTYKGGPLFPADI 60  
QY 61 LSPQOYSGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQOYSGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKKTGMWEASTERYPRDGLKGEIHKALKDKGHHYVEFKSI 180  
DB 121 KYKFIGVNFPSDGPVMQKKTGMWEASTERYPRDGLKGDIMALRLBGGHHYVEFKSI 180  
QY 181 YMAKPPVQLPGYVYVDSKLDITSHNEDYTIYQYERTEGRHHLFL 225  
DB 181 YMAKPPVQLPGYVYVDSKLDITSHNEDYTIYQYERTEGRHHLFL 225

DB 181 YMAKPPVQLPGYVYVDSKLDITSHNEDYTIYQYERTEGRHHLFL 225

## RESULT 53

US-10-931-304-81  
Sequence 81, Application US/10931304  
Publication No. US20050196768A1

## GENERAL INFORMATION:

APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert  
APPLICANT: Baird, Geoffrey  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
FILE REFERENCE: 39754-0831CP2CP3  
CURRENT FILING DATE: US/10/931,304  
CURRENT FILING DATE: 2004-08-30  
PRIOR APPLICATION NUMBER: 10/209,208  
PRIOR FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 81  
LENGTH: 234  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Dimer polypeptide variant "dimer2,2MM (dimer3) (dtomato)"  
US-10-931-304-81

Query Match 90.2%; Score 1095; DB 5; Length 234;  
Best Local Similarity 91.0%; Pred. No. 7e-102;  
Matches 201; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 5 KNVKEKFRFKVMEGTVNGHBEFIEGEGRPYEGHNTVTKYTKGGPLFPADIISPO 64  
DB 6 BEVKEKFRFKVMEGTVNGHBEFIEGEGRPYEGHNTVTKYTKGGPLFPADIISPO 65  
QY 65 FOYSGKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTYKVF 124  
DB 66 FMYSKAYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGTILYKXKM 125  
QY 125 IGVPSPDGPVMQKKTGMWEASTERYPRDGLKGEIHKALKDKGHHYVEFKSIYMAK 184  
DB 126 RGTNFPDGPVMQKKTGMWEASTERYPRDGLKGEIHKALKDKGHHYVEFKTIYMAK 185  
QY 185 KPVOQLPGYVYVDSKLDITSHNEDYTIYQYERTEGRHHLFL 225  
DB 186 KPVOQLPGYVYVDSKLDITSHNEDYTIYQYERTEGRHHLFL 226

## RESULT 54

US-10-931-304-106  
Sequence 106, Application US/10931304  
Publication No. US20050196768A1

## GENERAL INFORMATION:

APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert  
APPLICANT: Baird, Geoffrey  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
FILE REFERENCE: 39754-0831CP2CP3  
CURRENT FILING DATE: US/10/931,304  
CURRENT FILING DATE: 2004-08-30  
PRIOR APPLICATION NUMBER: 10/209,208  
PRIOR FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538



APPLICANT: Perez, Carl  
APPLICANT: Lindenbaum, Michael  
APPLICANT: Greene, Amy  
APPLICANT: Leung, Josephine  
APPLICANT: Fleming, Elena  
APPLICANT: Stewart, Sandra  
APPLICANT: Shellard, Joan  
TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
FILE REFERENCE: 17084-022002/420B  
CURRENT APPLICATION NUMBER: US/11/006,076  
CURRENT FILING DATE: 2004-12-06  
PRIOR APPLICATION NUMBER: 60/294,758  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/366,891  
PRIOR FILING DATE: 2002-03-21  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 40  
LENGTH: 230  
TYPE: PRT  
ORGANISM: Discosoma species  
US-11-006-076-40

Query Match 89.4%; Score 1085.5; DB 6; Length 230;  
Best Local Similarity 88.1%; Pred. No. 6.2e-101;  
Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;

QY 1 MRSSKNVKEFRMRKVRMEGTNGHGFIEBEGGRPYEGHNTYKLTQTKGGLPFPAADI 60  
DB 1 MSCSKNVKEFRMRKVRMEGTNGHGFIEBEGGRPYEGHNTYKLTQTKGGLPFPAADI 60  
QY 61 LSPQFQSSKYYVKGHPADIPYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQSSKYYVKGHPADIPYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
DB 121 KYKFIGVNFPSDGPVMQKTMGWEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHFL 226

RESULT 58  
US-10-006-922-46

Sequence 46, Application US/10006922  
Publication No. US20020197676A1

GENERAL INFORMATION:

APPLICANT: Lukyanov, Sergey A.  
APPLICANT: Prtakov, Arcady F.

APPLICANT: Labas, Yulii A.  
APPLICANT: Matz, Mikhail V.

APPLICANT: Terekikh, Alexey  
TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and  
TITLE OF INVENTION: Methods for Using the Same

FILE REFERENCE: CLON-035C1P

CURRENT APPLICATION NUMBER: US/10/006,922  
CURRENT FILING DATE: 2001-12-04

PRIOR APPLICATION NUMBER: 09/120,330  
PRIOR FILING DATE: 1998-12-11

PRIOR APPLICATION NUMBER: 09/457,898  
PRIOR FILING DATE: 1999-12-09

PRIOR APPLICATION NUMBER: 09/458,144  
PRIOR FILING DATE: 1999-12-09

PRIOR APPLICATION NUMBER: 09/458,477  
PRIOR FILING DATE: 1999-12-09

PRIOR APPLICATION NUMBER: 09/457,556  
PRIOR FILING DATE: 1999-12-09

PRIOR APPLICATION NUMBER: 09/444,338  
PRIOR FILING DATE: 1999-11-19

NUMBER OF SEQ ID NOS: 46

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 46  
LENGTH: 205  
TYPE: PRT  
ORGANISM: Discosoma species  
US-10-006-922-46

Query Match 86.2%; Score 1046; DB 4; Length 205;  
Best Local Similarity 86.6%; Pred. No. 5.1e-97;  
Matches 197; Conservative 3; Mismatches 5; Indels 20; Gaps 1;

QY 1 MRSSKNVKEFRMRKVRMEGTNGHGFIEBEGGRPYEGHNTYKLTQTKGGLPFPAADI 60  
DB 1 MRSSKNVKEFRMRKVRMEGTNGHGFIEBEGGRPYEGHNTYKLTQTKGGLPFPAADI 60  
QY 61 LSPQFQSSKYYVKGHPADIPYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQSSKYYVKGHPADIPYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
DB 121 KYKFIGVNFPSDGPVMQKTMGWEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 160  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHFL 205

RESULT 59  
US-10-121-258-8

Sequence 8, Application US/10121258  
Publication No. US20030059835A1

GENERAL INFORMATION:

APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert

TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
FILE REFERENCE: UC083.1CR2CPL

CURRENT APPLICATION NUMBER: US/10/121,258  
CURRENT FILING DATE: 2002-04-10

PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26

PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24

NUMBER OF SEQ ID NOS: 78

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8

LENGTH: 225

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Dered polypeptide variant "mRFP1"

US-10-121-258-8

Query Match 84.1%; Score 1021; DB 4; Length 225;  
Best Local Similarity 86.9%; Pred. No. 1.9e-94;  
Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFRMRKVRMEGTNGHGFIEBEGGRPYEGHNTYKLTQTKGGLPFPAADI 60  
DB 1 MRSSKNVKEFRMRKVRMEGTNGHGFIEBEGGRPYEGHNTYKLTQTKGGLPFPAADI 60  
QY 61 LSPQFQSSKYYVKGHPADIPYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQSSKYYVKGHPADIPYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
DB 121 KYKFIGVNFPSDGPVMQKTMGWEASTERTLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRH 221  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRH 221



Oy 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRH 221  
Db 181 YMAKPPVOLPGAYKLDYKLDITSHNEDYTIYEQYERABGRH 221

RESULT 63  
US-10-931-304-79

/ Sequence 79, Application US/10931304  
/ Publication No. US20050196768A1

/ GENERAL INFORMATION:

/ APPLICANT: Tsien, Roger

/ APPLICANT: Campbell, Robert

/ APPLICANT: Baird, Geoffrey

/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT

/ FILE REFERENCE: 39754-0831CP2C3

/ CURRENT FILING DATE: 2004-08-30

/ PRIOR FILING DATE: 2004-08-30

/ PRIOR FILING DATE: 2002-07-29

/ PRIOR FILING DATE: 2002-04-10

/ PRIOR FILING DATE: 2001-05-24

/ PRIOR FILING DATE: 2001-02-26

/ NUMBER OF SEQ ID NOS: 110

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 79

/ LENGTH: 225

/ TYPE: PRT

/ ORGANISM: Artificial Sequence

/ FEATURE:

/ OTHER INFORMATION: Dered polypeptide variant "mRFP1.1"

US-10-931-304-79

Query Match 83.4%; Score 1012; DB 5; Length 225;

Best Local Similarity 86.0%; Pred. No. 1.6e-93;

Matches 190; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

Oy 1 MRSSKNVKEFMRFVMEGVNGHFEIEBEGGRPYEGHNTVYKLTGKGGLPFANDI 60

Db 1 MASSEDVKEFMRFVMEGVNGHFEIEBEGGRPYEGHNTVYKLTGKGGLPFANDI 60

Oy 61 LSPQFOYGSKYVYVHPADIPYKLSPEEGFKMERVNNFEDGGVTVTQDSSLQDGEFIY 120

Db 61 LSPQFMYSKAYVYVHPADIPYKLSPEEGFKMERVNNFEDGGVTVTQDSSLQDGEFIY 120

Oy 121 KVFRTGVNPFSDGVPWQKTMGWEASTERLVPKGLKGEIHKALKDKDGGHYLVEFCSI 180

Db 121 KVKRGTFNPFSDGVPWQKTMGWEASTERLVPKGLKGEIHKALKDKDGGHYLVEFCKT 180

Oy 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRH 221

Db 181 YMAKPPVOLPGAYKLDYKLDITSHNEDYTIYEQYERABGRH 221

RESULT 64  
US-10-931-304-104

/ Sequence 104, Application US/10931304

/ Publication No. US20050196768A1

/ GENERAL INFORMATION:

/ APPLICANT: Tsien, Roger

/ APPLICANT: Campbell, Robert

/ APPLICANT: Baird, Geoffrey

/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT

/ FILE REFERENCE: 39754-0831CP2C3

/ CURRENT FILING DATE: 2004-08-30

/ PRIOR FILING DATE: 2004-08-30

/ PRIOR FILING DATE: 2002-07-29

/ NUMBER OF SEQ ID NOS: 110

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 100

/ LENGTH: 225

/ TYPE: PRT

/ ORGANISM: Artificial Sequence

/ FEATURE:

/ OTHER INFORMATION: Dered polypeptide variant "mTangerine"

US-10-931-304-104

Query Match 82.5%; Score 1002; DB 5; Length 225;

Best Local Similarity 85.5%; Pred. No. 1.6e-92;

Matches 189; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

Oy 1 MRSSKNVKEFMRFVMEGVNGHFEIEBEGGRPYEGHNTVYKLTGKGGLPFANDI 60

Db 1 MASSEDVKEFMRFVMEGVNGHFEIEBEGGRPYEGHNTVYKLTGKGGLPFANDI 60

Oy 61 LSPQFOYGSKYVYVHPADIPYKLSPEEGFKMERVNNFEDGGVTVTQDSSLQDGEFIY 120

Db 61 LSPQFMYSKAYVYVHPADIPYKLSPEEGFKMERVNNFEDGGVTVTQDSSLQDGEFIY 120

Oy 121 KVFRTGVNPFSDGVPWQKTMGWEASTERLVPKGLKGEIHKALKDKDGGHYLVEFCSI 180

Db 121 KVKRGTFNPFSDGVPWQKTMGWEASTERLVPKGLKGEIHKALKDKDGGHYLVEFCKT 180

Oy 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRH 221

Db 181 YMAKPPVOLPGAYKLDYKLDITSHNEDYTIYEQYERABGRH 221

/ PRIOR APPLICATION NUMBER: 10/121,258  
/ PRIOR FILING DATE: 2002-04-10  
/ PRIOR APPLICATION NUMBER: 09/866,538  
/ PRIOR FILING DATE: 2001-05-24  
/ PRIOR APPLICATION NUMBER: 09/794,308  
/ PRIOR FILING DATE: 2001-02-26  
/ NUMBER OF SEQ ID NOS: 110  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 104  
/ LENGTH: 225  
/ TYPE: PRT  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Dered polypeptide variant "mHoneydew"

US-10-931-304-104

Query Match 82.9%; Score 1007; DB 5; Length 225;

Best Local Similarity 85.1%; Pred. No. 5e-93;

Matches 189; Conservative 8; Mismatches 25; Indels 0; Gaps 0;

Oy 1 MRSSKNVKEFMRFVMEGVNGHFEIEBEGGRPYEGHNTVYKLTGKGGLPFANDI 60

Db 1 MASSEDVKEFMRFVMEGVNGHFEIEBEGGRPYEGHNTVYKLTGKGGLPFANDI 60

Oy 61 LSPQFOYGSKYVYVHPADIPYKLSPEEGFKMERVNNFEDGGVTVTQDSSLQDGEFIY 120

Db 61 LSPQFMYSKAYVYVHPADIPYKLSPEEGFKMERVNNFEDGGVTVTQDSSLQDGEFIY 120

Oy 121 KVFRTGVNPFSDGVPWQKTMGWEASTERLVPKGLKGEIHKALKDKDGGHYLVEFCSI 180

Db 121 KVKRGTFNPFSDGVPWQKTMGWEASTERLVPKGLKGEIHKALKDKDGGHYLVEFCKT 180

Oy 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRH 221

Db 181 YMAKPPVOLPGAYKLDYKLDITSHNEDYTIYEQYERABGRH 221

RESULT 65  
US-10-931-304-100

/ Sequence 100, Application US/10931304

/ Publication No. US20050196768A1

/ GENERAL INFORMATION:

/ APPLICANT: Tsien, Roger

/ APPLICANT: Campbell, Robert

/ APPLICANT: Baird, Geoffrey

/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT

/ FILE REFERENCE: 39754-0831CP2C3

/ CURRENT FILING DATE: 2004-08-30

/ PRIOR FILING DATE: 2004-08-30

/ PRIOR FILING DATE: 2002-07-29

/ PRIOR FILING DATE: 2002-04-10

/ PRIOR FILING DATE: 2001-05-24

/ PRIOR FILING DATE: 2001-02-26

/ NUMBER OF SEQ ID NOS: 110

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 100

/ LENGTH: 225

/ TYPE: PRT

/ ORGANISM: Artificial Sequence

/ FEATURE:

/ OTHER INFORMATION: Dered polypeptide variant "mTangerine"

US-10-931-304-100

Query Match 82.5%; Score 1002; DB 5; Length 225;

Best Local Similarity 85.5%; Pred. No. 1.6e-92;

Matches 189; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

Oy 1 MRSSKNVKEFMRFVMEGVNGHFEIEBEGGRPYEGHNTVYKLTGKGGLPFANDI 60

Db 1 MASSEDVKEFMRFVMEGVNGHFEIEBEGGRPYEGHNTVYKLTGKGGLPFANDI 60

Oy 61 LSPQFOYGSKYVYVHPADIPYKLSPEEGFKMERVNNFEDGGVTVTQDSSLQDGEFIY 120

Db 61 LSPQFMYSKAYVYVHPADIPYKLSPEEGFKMERVNNFEDGGVTVTQDSSLQDGEFIY 120

Oy 121 KVFRTGVNPFSDGVPWQKTMGWEASTERLVPKGLKGEIHKALKDKDGGHYLVEFCSI 180

Db 121 KVKRGTFNPFSDGVPWQKTMGWEASTERLVPKGLKGEIHKALKDKDGGHYLVEFCKT 180

Oy 181 YMAKPPVOLPGYYVDSKLDITSHNEDYTIYEQYERTGRH 221

Db 181 YMAKPPVOLPGAYKLDYKLDITSHNEDYTIYEQYERABGRH 221

```

Db      1  MASSSDVYKEMRRFVNRMEGSGVNGHEFEIISGEGRPYEGQTALKYTKGGPLPFAMDI 60
Qy      61  LSPQGVSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY 120
Db      61  LSPQGVSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGERITY 120
Qy      121  KYKPIGVNFPDGPVMQKKTGMWEASSTERLYPRDGLKGEIHKALKDGGHYLVFYSI 180
Db      121  KYKLRGTNFPDGPVMQKKTGMWEASSTERLYPRDGLKGEIHKALKDGGHYDAEVKTT 180
Qy      181  YMAKKEVQLPGAYKVDIKLDTISHNEDYTIIVEQYERTEGRH 221
Db      181  YMAKKEVQLPGAYKVDIKLDTISHNEDYTIIVEQYERTEGRH 221

```

## RESULT 66

```

US-10-931-304-85
; Sequence 85, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baile, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2C3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 85
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dared polypeptide variant "Ors4-9"
US-10-931-304-85

```

```

Query Match      81.9%; Score 994; DB 5; Length 236;
Best Local Similarity 86.5%; Pred. No. 1.1e-91;
Matches 186; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

```

```

Qy      7  VIKERFRKVRMEGTNGHFEIISGEGRPYEGHNTVTKLYTKGGPLPFAMDISPQFO 66
Db      12  VIKERFRKVRMEGTNGHFEIISGEGRPYEGHNTVTKLYTKGGPLPFAMDISPQFT 71
Qy      67  YGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTYKXFTG 126
Db      72  YGSKAYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGEFTYKVLHG 131
Qy      127  VNFPSDGPVMQKKTGMWEASSTERLYPRDGLKGEIHKALKDGGHYLVFYSIYMAKKP 186
Db      132  TNFPSDGPVMQKKTGMWEASSTERLYPRDGLKGEIHKALKDGGHYDAEVKTTYMAKKP 191
Qy      187  VOLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRH 221
Db      192  VOLPGAYKVDIKLDTISHNEDYTIIVEQYERTEGRH 226

```

## RESULT 67

```

US-10-931-304-89
; Sequence 89, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baile, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2C3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dared polypeptide variant "mFRFP (F2Q6) (mGrape2)"
US-10-931-304-89

```

```

Query Match      81.7%; Score 992; DB 5; Length 236;
Best Local Similarity 86.0%; Pred. No. 1.7e-91;
Matches 185; Conservative 6; Mismatches 24; Indels 0; Gaps 0;

```

```

Qy      7  VIKERFRKVRMEGTNGHFEIISGEGRPYEGHNTVTKLYTKGGPLPFAMDISPQFO 66
Db      12  VIKERFRKVRMEGTNGHFEIISGEGRPYEGHNTVTKLYTKGGPLPFAMDISPQFM 71
Qy      67  YGSKYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTYKXFTG 126
Db      72  YGSKAYVYKHPADIPDYKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGEFTYKVLHG 131
Qy      127  VNFPSDGPVMQKKTGMWEASSTERLYPRDGLKGEIHKALKDGGHYLVFYSIYMAKKP 186
Db      132  TNFPSDGPVMQKKTGMWEASSTERLYPRDGLKGEIHKALKDGGHYDAEVKTTYMAKKP 191
Qy      187  VOLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRH 221
Db      192  VOLPGAYKVDIKLDTISHNEDYTIIVEQYERTEGRH 226

```

## RESULT 68

```

US-10-931-304-102
; Sequence 102, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baile, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2C3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102
; LENGTH: 236
; TYPE: PRT

```





```

; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dated polypeptide variant "mYOP (74-11)"
; NAME/KEY: VARIANT
; LOCATION: 62
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-931-304-94
```

```
Query Match      81.2%; Score 986; DB 5; Length 235;
Best Local Similarity 85.6%; Pred. No. 7e-91;
Matches 184; Conservative 5; Mismatches 26; Indels 0; Gaps 0;
```

```
QY 7 VIKEMRRFVYRMGEVTVNGHFEFIEGEGRPYEGHNTVTLKTKGGLPFPAMDILSPQFG 66
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 12 IIEKFRFVYRMGEVTVNGHFEFIEGEGRPYEGHNTVTLKTKGGLPFPAMDILSPQFG 71
QY 67 YGSKVYVYKHPADIPDYKLSFPEGFKMERVMNFDGVTVTQDSSLQDGCFTYKVKFG 126
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 72 YGSKAVYKHPADIPDYKLSFPEGFKMERVMNFDGVTVTQDSSLQDGCFTYKVKLRG 131
QY 127 VNFPSDGPVWQKKTWMEASSTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 132 TNFPSDGPVWQKKTWMEASSTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 191
QY 187 VOLPGYIVYDSKLDITSHNEDYTIYQYERTEGRH 221
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 192 VOLPGAYIVGIKLDITSHNEDYTIYQYERTEGRH 226
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 72
US-10-931-304-96
; Sequence 96, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dated polypeptide variant "mRPP (A2/6-6)"
US-10-931-304-96
```

```
Query Match      80.1%; Score 973; DB 5; Length 236;
Best Local Similarity 84.7%; Pred. No. 1.4e-89;
```

```
Matches 182; Conservative 6; Mismatches 27; Indels 0; Gaps 0;
QY 7 VIKEMRRFVYRMGEVTVNGHFEFIEGEGRPYEGHNTVTLKTKGGLPFPAMDILSPQFG 66
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 12 IIEKFRFVYRMGEVTVNGHFEFIEGEGRPYEGHNTVTLKTKGGLPFPAMDILSPQFG 71
QY 67 YGSKVYVYKHPADIPDYKLSFPEGFKMERVMNFDGVTVTQDSSLQDGCFTYKVKFG 126
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 72 YGSKAVYKHPADIPDYKLSFPEGFKMERVMNFDGVTVTQDSSLQDGCFTYKVKLRG 131
QY 127 VNFPSDGPVWQKKTWMEASSTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 132 TNFPSDGPVWQKKTWMEASSTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 191
QY 187 VOLPGYIVYDSKLDITSHNEDYTIYQYERTEGRH 221
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 192 VOLPGAYIVGIKLDITSHNEDYTIYQYERTEGRH 226
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 73
US-10-931-304-98
; Sequence 98, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dated polypeptide variant "mStrawberry"
US-10-931-304-98
```

```
Query Match      80.1%; Score 973; DB 5; Length 236;
Best Local Similarity 84.7%; Pred. No. 1.4e-89;
Matches 182; Conservative 6; Mismatches 27; Indels 0; Gaps 0;
```

```
QY 7 VIKEMRRFVYRMGEVTVNGHFEFIEGEGRPYEGHNTVTLKTKGGLPFPAMDILSPQFG 66
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 12 IIEKFRFVYRMGEVTVNGHFEFIEGEGRPYEGHNTVTLKTKGGLPFPAMDILSPQFG 71
QY 67 YGSKVYVYKHPADIPDYKLSFPEGFKMERVMNFDGVTVTQDSSLQDGCFTYKVKFG 126
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 72 YGSKAVYKHPADIPDYKLSFPEGFKMERVMNFDGVTVTQDSSLQDGCFTYKVKLRG 131
QY 127 VNFPSDGPVWQKKTWMEASSTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 132 TNFPSDGPVWQKKTWMEASSTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 191
QY 187 VOLPGYIVYDSKLDITSHNEDYTIYQYERTEGRH 221
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 192 VOLPGAYIVGIKLDITSHNEDYTIYQYERTEGRH 226
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
RESULT 74
US-10-931-304-87
```

```
; Sequence 87, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dated polypeptide variant "Y1.3 (mYOF 1.3) (mBanana)"
US-10-931-304-87

Query Match          79.1%; Score 960; DB 5; Length 236;
Best Local Similarity 83.7%; Pred. No. 2.9e-88;
Matches 180; Conservative 5; Mismatches 30; Indels 0; Gaps 0;

QY 7 VIKERFRKVMNEGTVNGHEFIEBEGEGRPYEGHNTVKLKYTKGGLPFPAMDILSPQFC 66
    |||||
    12 VIKERFRKVMNEGTVNGHEFIEBEGEGRPYEGTQAKLKYTKGGLPFPAMDILSPQFC 71
    |||||

QY 67 YGSKVYVGHPRADIPDYKLSFPPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTYVKRFIG 126
    |||||
    72 YGSKAVYVGHPRADIPDYKLSFPPEGFKMERVMNPFEDGVTVAQDSSLQDGEFTYVKRLRG 131
    |||||

QY 127 VNFPSDGVPMQKKTWGMGAEATERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186
    |||||
    132 TNFSDGVPMQKKTWGMGAEATERLYPRDGVLKGEIHKALKLKDGGHYSAETKTYKAKKP 191
    |||||

QY 187 VOLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRH 221
    |||||
    192 VOLPGAVYVAGEKIDITSHNEDYTIIVELYERAEGRH 226
    |||||

RESULT 75
US-10-724-178-1041
; Sequence 1041, Application US/10724178
; Publication No. US20040137528A1
; GENERAL INFORMATION:
; APPLICANT: Odysee Therapeutics, Inc.
; APPLICANT: Michnick, Stephen
; APPLICANT: Macdonald, Marnie
; APPLICANT: Lamerdin, Jane
; TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT
; FILE REFERENCE: ODDY007
; CURRENT APPLICATION NUMBER: US/10/724,178
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US 60/461,133
; PRIOR FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1041
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
```

```
; OTHER INFORMATION: modif. frag.; RFP FLF, with position 1 Met removed
US-10-724-178-1041

Query Match          69.9%; Score 848; DB 4; Length 183;
Best Local Similarity 87.4%; Pred. No. 4.1e-77;
Matches 159; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 3 SSKVIVKFMRFKVMNEGTVNGHEFIEBEGEGRPYEGHNTVKLKYTKGGLPFPAMDILS 62
    |||||
    2 SSEDVIRKFMRFKVMNEGTVNGHEFIEBEGEGRPYEGTQAKLKYTKGGLPFPAMDILS 61
    |||||

QY 63 PQFGYGSKAVYVGHPRADIPDYKLSFPPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTYVK 122
    |||||
    62 PQFGYGSKAVYVGHPRADIPDYKLSFPPEGFKMERVMNPFEDGVTVTQDSSLQDGEFTYKV 121
    |||||

QY 123 KFTGVNFPDGVPMQKKTWGMGAEATERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSIYM 182
    |||||
    122 KLRGTNFPDGVPMQKKTWGMGAEATERLYPRDGVLKGEIHKALKLKDGGHYDAEVKTYM 181
    |||||

QY 183 AK 184
    ||
    182 AK 183
    ||
```

Search completed: January 11, 2006, 02:12:06  
Job time : 64 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using sw model

Run on: January 11, 2006, 02:08:34 ; Search time 9 Seconds  
(without alignments)  
212.341 Million cell updates/sec

Title: US-10-006-922A-12  
Sequence: 1214  
1 MRSSKNVKEFMRFKVRMEG.....EDYTVIEQYERTGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 61141 seqs, 8493638 residues

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Database : Published Applications AA New:  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	100.0	225	6	US-10-209-208-1 Sequence 1, Appli
2	1207	99.4	225	6	US-10-209-208-20 Sequence 20, Appli
3	1160	95.6	225	6	US-10-209-208-4 Sequence 4, Appli
4	1153	95.0	225	6	US-10-209-208-24 Sequence 24, Appli
5	1121	92.3	226	6	US-10-209-208-6 Sequence 6, Appli
6	1021	84.1	225	6	US-10-209-208-8 Sequence 8, Appli
7	1012	83.4	225	6	US-10-209-208-79 Sequence 79, Appli

## ALIGNMENTS

RESULT 1  
US-10-209-208-1  
Sequence 1, Application US/10209208  
Publication No. US20050244921A1  
GENERAL INFORMATION:  
APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert  
APPLICANT: Geoffrey Baird  
TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
TITLE OF INVENTION: FOR MAKING SAME  
FILE REFERENCE: UC083.1CP2CP2  
CURRENT APPLICATION NUMBER: US/10/209,208  
CURRENT FILING DATE: 2002-07-29

PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Dicosoma sp.  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(225)  
OTHER INFORMATION: wild-type Dared  
US-10-209-208-1

Query Match 100.0%; Score 1214; DB 6; Length 225;  
Best Local Similarity 100.0%; Pred. No. 2.3e-110; Indels 0; Gaps 0;  
Matches 225; Conservative 0; Mismatches 0;

QY	1	MRSSKNVKEFMRFKVRMEGTGNGHEFEIBGBGRPYEGHNTVTKVTGKGPFPAMD	60
DB	1	MRSSKNVKEFMRFKVRMEGTGNGHEFEIBGBGRPYEGHNTVTKVTGKGPFPAMD	60
QY	61	LSPOFQYSKYVYKHPADIPYKLSPEGRKWERVNFEDGVTYVQDSLODGCFTY	120
DB	61	LSPOFQYSKYVYKHPADIPYKLSPEGRKWERVNFEDGVTYVQDSLODGCFTY	120
QY	121	KYKFGVNPSPDGPVMOCKTGWMEASTERLVPDGLGGEIHKALKLDGGHYLVEFKSI	180
DB	121	KYKFGVNPSPDGPVMOCKTGWMEASTERLVPDGLGGEIHKALKLDGGHYLVEFKSI	180
QY	181	YMAKRPVQLPGYVYVDSKLDITSNEDYTVIEQYERTGRHHLFL 225	
DB	181	YMAKRPVQLPGYVYVDSKLDITSNEDYTVIEQYERTGRHHLFL 225	

RESULT 2  
US-10-209-208-20  
Sequence 20, Application US/10209208  
Publication No. US20050244921A1  
GENERAL INFORMATION:  
APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert  
APPLICANT: Geoffrey Baird  
TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
TITLE OF INVENTION: FOR MAKING SAME  
FILE REFERENCE: UC083.1CP2CP2  
CURRENT APPLICATION NUMBER: US/10/209,208  
CURRENT FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 20  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Dared with 1125R  
US-10-209-208-20

Query Match 99.4%; Score 1207; DB 6; Length 225;  
Best Local Similarity 99.6%; Pred. No. 1.1e-109; Indels 0; Gaps 0;  
Matches 224; Conservative 0; Mismatches 1;

QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIBGBGRPYEGHNTVTKVTGKGPFPAMD 60

```

Db      1 MRSSKNVYKEFMRFKVRMEGVNNGHEFEIEEGEGRPFEGHNTVTKLTKYTKGGPLPFANDI 60
        |||
Qy      61 LSPQFQYSKYVVKHPADIPDYKLSFPEGFKMERVNNFEDGGVTVTVQDSSLQDGCFTY 120
        |||
Db      61 LSPQFQYSKYVVKHPADIPDYKLSFPEGFKMERVNNFEDGGVTVTVQDSSLQDGCFTY 120
        |||
Qy      121 KVKFIGNVPFSDGPMQKKTGMWEASTERLYPRDGLVKGELHKAALKDGGHYLVEFKSI 180
        |||
Db      121 KVKFIGNVPFSDGPMQKKTGMWEASTERLYPRDGLVKGELHKAALKDGGHYLVEFKSI 180
        |||
Qy      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||
Db      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||

```

## RESULT 3

```

US-10-209-208-4
; Sequence 4, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dered polypeptide variant "T1"
US-10-209-208-4

```

```

Query Match      95.6%; Score 1160; DB 6; Length 225;
Best Local Similarity 96.0%; Pred. No. 3.7e-105;
Matches 216; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy      1 MRSSKNVYKEFMRFKVRMEGVNNGHEFEIEEGEGRPFEGHNTVTKLTKYTKGGPLPFANDI 60
        |||
Db      1 MASSSDVYKEFMRFKVRMEGVNNGHEFEIEEGEGRPFEGHNTVTKLTKYTKGGPLPFANDI 60
        |||
Qy      61 LSPQFQYSKYVVKHPADIPDYKLSFPEGFKMERVNNFEDGGVTVTVQDSSLQDGCFTY 120
        |||
Db      61 LSPQFQYSKYVVKHPADIPDYKLSFPEGFKMERVNNFEDGGVTVTVQDSSLQDGCFTY 120
        |||
Qy      121 KVKFIGNVPFSDGPMQKKTGMWEASTERLYPRDGLVKGELHKAALKDGGHYLVEFKSI 180
        |||
Db      121 KVKFIGNVPFSDGPMQKKTGMWEASTERLYPRDGLVKGELHKAALKDGGHYLVEFKSI 180
        |||
Qy      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||
Db      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||

```

## RESULT 4

```

US-10-209-208-24
; Sequence 24, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert

```

```

; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dered polypeptide variant "T1" with 1125R mutation
US-10-209-208-24

```

```

Query Match      95.0%; Score 1153; DB 6; Length 225;
Best Local Similarity 95.6%; Pred. No. 1.8e-104;
Matches 215; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      1 MRSSKNVYKEFMRFKVRMEGVNNGHEFEIEEGEGRPFEGHNTVTKLTKYTKGGPLPFANDI 60
        |||
Db      1 MASSSDVYKEFMRFKVRMEGVNNGHEFEIEEGEGRPFEGHNTVTKLTKYTKGGPLPFANDI 60
        |||
Qy      61 LSPQFQYSKYVVKHPADIPDYKLSFPEGFKMERVNNFEDGGVTVTVQDSSLQDGCFTY 120
        |||
Db      61 LSPQFQYSKYVVKHPADIPDYKLSFPEGFKMERVNNFEDGGVTVTVQDSSLQDGCFTY 120
        |||
Qy      121 KVKFIGNVPFSDGPMQKKTGMWEASTERLYPRDGLVKGELHKAALKDGGHYLVEFKSI 180
        |||
Db      121 KVKFIGNVPFSDGPMQKKTGMWEASTERLYPRDGLVKGELHKAALKDGGHYLVEFKSI 180
        |||
Qy      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||
Db      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||

```

## RESULT 5

```

US-10-209-208-6
; Sequence 6, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dered polypeptide variant "dimer2"
US-10-209-208-6

```

```

Query Match      92.3%; Score 1121; DB 6; Length 226;

```

Best Local Similarity 92.8%; Pred. No. 2.2e-101;  
Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 3 SSKNVKEFMFKVMEGTNGHEFEIEGEGRGPRYEGHNTVKLKVTGGLPFPAMDILS 62  
DB 4 SSEDVKEFMFKVMEGTVNGHEFEIEGEGRGPRYEGHNTVKLKVTGGLPFPAMDILS 63  
QY 63 POFQYGSKVVYVHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTODSSLQDGCFTYK 122  
DB 64 POFQYGSKVVYVHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTODSSLQDGCFTYK 123  
QY 123 KPIGVNFPSPDGPVWQKKTGMEASTERLYPRDGVKGEIHKALKKDGSHYLVFKSIYM 182  
DB 124 KPIGVNFPSPDGPVWQKKTGMEASTERLYPRDGVKGEIHKALKKDGSHYLVFKSIYM 183  
QY 183 AKKPVQLPGYVYVDSKLDITSHNEDYTIYEQYERABGRH 225  
DB 184 AKKPVQLPGYVYVDSKLDITSHNEDYTIYEQYERABGRH 226

RESULT 6  
US-10-209-208-8  
Sequence 8, Application US/10209208  
Publication No. US20050244921A1  
GENERAL INFORMATION:  
APPLICANT: Tsiem, Roger  
APPLICANT: Campbell, Robert  
APPLICANT: Geoffrey, Baird  
TITLE OF INVENTION: FOR MAKING SAME  
FILE REFERENCE: UC083.1CP2CP2  
CURRENT APPLICATION NUMBER: US/10/209,208  
CURRENT FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: FaastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Dated polypeptide variant "mrPp1"

US-10-209-208-8  
Query Match 84.1%; Score 1021; DB 6; Length 225;  
Best Local Similarity 86.9%; Pred. No. 1e-91;  
Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMFKVMEGTNGHEFEIEGEGRGPRYEGHNTVKLKVTGGLPFPAMDILS 60  
DB 1 MSSSDVKEFMFKVMEGTVNGHEFEIEGEGRGPRYEGHNTVKLKVTGGLPFPAMDILS 60  
QY 61 LSPQYGSKVVYVHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTODSSLQDGCFTY 120  
DB 61 LSPQYGSKVVYVHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTODSSLQDGCFTY 120  
QY 121 KYKFGVNFPSDGPVWQKKTGMEASTERLYPRDGVKGEIHKALKKDGSHYLVFKSIYM 180  
DB 121 KYKFGVNFPSDGPVWQKKTGMEASTERLYPRDGVKGEIHKALKKDGSHYLVFKSIYM 180  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIYEQYERABGRH 221  
DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIYEQYERABGRH 221

RESULT 7  
US-10-209-208-79  
Sequence 79, Application US/10209208

Publication No. US20050244921A1  
GENERAL INFORMATION:  
APPLICANT: Tsiem, Roger  
APPLICANT: Campbell, Robert  
APPLICANT: Geoffrey, Baird  
TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
TITLE OF INVENTION: FOR MAKING SAME  
FILE REFERENCE: UC083.1CP2CP2  
CURRENT APPLICATION NUMBER: US/10/209,208  
CURRENT FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: FaastSeq for Windows Version 4.0  
SEQ ID NO 79  
LENGTH: 225  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Dated polypeptide variant "mrPp1.1"

US-10-209-208-79  
Query Match 83.4%; Score 1012; DB 6; Length 225;  
Best Local Similarity 86.0%; Pred. No. 7.4e-91;  
Matches 190; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 1 MSSKNVKEFMFKVMEGTNGHEFEIEGEGRGPRYEGHNTVKLKVTGGLPFPAMDILS 60  
DB 1 MSSSDVKEFMFKVMEGTVNGHEFEIEGEGRGPRYEGHNTVKLKVTGGLPFPAMDILS 60  
QY 61 LSPQYGSKVVYVHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTODSSLQDGCFTY 120  
DB 61 LSPQYGSKVVYVHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTODSSLQDGCFTY 120  
QY 121 KYKFGVNFPSDGPVWQKKTGMEASTERLYPRDGVKGEIHKALKKDGSHYLVFKSIYM 180  
DB 121 KYKFGVNFPSDGPVWQKKTGMEASTERLYPRDGVKGEIHKALKKDGSHYLVFKSIYM 180  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIYEQYERABGRH 221  
DB 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIYEQYERABGRH 221

Search completed: January 11, 2006, 02:12:24  
Job time : 10 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_p2n model

Run on: January 12, 2006, 12:23:38 ; Search time 3690 Seconds  
(without alignments)  
3466.064 Million cell updates/sec

Title: US-10-006-922A-12

Perfect score: 1214  
Sequence: 1 MRSSKNVKEPMRFKVMG.....EDYIVGEYERTEGRHHLFL 225

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5683141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 85

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:

MODEL=frame+22n.model -DEV=xlj  
-O=/cgn2.1/USPTO\_epoc1\_p/US10006222/runat\_10012006.162504.19114/app\_query.fasta\_1.391  
-DB=genEmbl -QFMT=fastap -SUFFIX=p2mlnscr.rge -MINMATCH=0.1 -LOOFCI=0  
-LIST=500 -DOCALIGN=200 -STRAT=1 -END=-1 -MATRIX=blowsum62 -TRANS=human40.cdi  
-MODE=LOCAL -OUTFMT=pcio -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922 @CGN 1.1 7415 @runat\_10012006.162504.19114 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGIOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*

1: gb\_ba:\*

2: gb\_in:\*

3: gb\_env:\*

4: gb\_cm:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pr:\*

9: gb\_ro:\*

10: gb\_scs:\*

11: gb\_ey:\*

12: gb\_un:\*

13: gb\_vl:\*

14: gb\_hcg:\*

15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	678	6	AR183915 Sequence
2	1214	100.0	678	6	AX172854 Sequence
3	1214	100.0	678	6	AX207715 Sequence

4	1214	100.0	678	6	AX233581 Sequence
5	1214	100.0	678	6	AX370404 Sequence
6	1214	100.0	678	6	AX824725 Sequence
7	1214	100.0	859	2	AF168419 Discosoma
8	1214	100.0	859	6	AR636082 Sequence
9	1214	100.0	859	6	AR670157 Sequence
10	1214	100.0	859	6	AX463698 Sequence
11	1214	100.0	681	6	AX233584 Sequence
12	1210	99.7	713	6	AX233627 Sequence
13	1210	99.7	723	6	AK527331 Sequence
14	1210	99.7	721	6	CQ882115 Sequence
15	1210	99.7	2772	6	CQ882117 Sequence
16	1210	99.7	4692	6	AX463702 Sequence
17	1210	99.7	6893	6	AX823860 Sequence
18	1210	99.7	8811	11	AY569779 Cloning v
19	1210	99.7	9320	6	AX663075 Sequence
20	1210	99.7	12404	11	AY569780 Cloning v
21	1207	99.4	678	6	AX370406 Sequence
22	1196	98.5	666	6	AX348043 Sequence
23	1196	98.5	666	6	AX353910 Sequence
24	1196	98.5	711	2	AY679107 Discosoma
25	1193	98.3	678	2	AF545828 Discosoma
26	1191	98.1	678	6	AX370408 Sequence
27	1191	98.1	678	6	AX824732 Sequence
28	1191	98.1	921	2	AY679106 Discosoma
29	1186	97.7	675	6	AX824731 Sequence
30	1186	97.7	1050	6	AX666133 Sequence
31	1186	97.7	4488	6	CQ849509 Sequence
32	1186	97.7	4546	11	AY818375 Cloning v
33	1186	97.7	4555	11	DQ005468 Cloning v
34	1186	97.7	4570	11	AY818373 Cloning v
35	1186	97.7	4829	11	AY640628 SiRNA vec
36	1186	97.7	4906	11	AY640625 SiRNA vec
37	1186	97.7	5311	11	AJ851284 Cloning v
38	1186	97.7	6423	11	AY613997 Cloning v
39	1186	97.7	6990	11	AY640634 SiRNA vec
40	1186	97.7	7058	11	AY640630 SiRNA vec
41	1186	97.7	7147	11	AY640630 SiRNA vec
42	1186	97.7	7616	6	CQ849511 Sequence
43	1186	97.7	9468	6	CS018252 Sequence
44	1186	97.7	10140	6	CS018253 Sequence
45	1186	97.7	10522	6	CS018254 Sequence
46	1186	97.7	10786	6	CS018256 Sequence
47	1186	97.7	11251	6	CS018251 Sequence
48	1186	97.7	11570	6	CS018250 Sequence
49	1186	97.7	11867	6	CS018255 Sequence
50	1186	97.7	11920	6	CS018249 Sequence
51	1186	97.7	14720	6	CS018257 Sequence
52	1186	97.7	16157	6	CQ981073 Sequence
53	1186	97.7	16157	6	CQ981074 Sequence
54	1184	97.5	898	6	AX686888 Sequence
55	1170	96.4	3441	11	AY916793 Dual flno
56	1170	96.4	5311	11	AJ851285 Cloning v
57	1170	96.4	13079	11	AJ851289 Cloning v
58	1165	96.0	5311	11	AJ851286 Cloning v
59	1165	96.0	10141	11	AY342347 Red H-Pel
60	1165	96.0	10276	11	AY342348 Red H-Pel
61	1165	96.0	10481	11	AY490568 URS- Red S
62	1165	96.0	13079	11	AJ851290 Cloning v
63	1121	92.3	681	11	AF506025 Synthetic
64	1121	92.3	1395	11	AF506026 Synthetic
65	1119	92.2	678	6	AX824729 Sequence
66	1095	90.2	705	11	AY678268 Synthetic
67	1095	90.2	1431	11	AY678269 Synthetic
68	1085.5	89.4	876	2	AF272711 Discosoma
69	1085.5	89.4	876	6	AX686894 Sequence
70	1021	84.1	678	11	AF506027 Synthetic
71	1021	84.1	696	11	AB166761 Synthetic
72	1021	84.1	4533	11	DQ005472 Cloning v
73	1021	84.1	4555	11	DQ005474 Cloning v
74	1021	84.1	4558	11	DQ005475 Cloning v
75	1021	84.1	5311	11	AJ851287 Cloning v
76	1021	84.1	13079	11	AJ851291 Cloning v

```
77 1007 82.9 678 11 AY678271 Synthetic
78 1002 82.5 678 11 AY678270 Synthetic
79 994 81.9 702 11 AY679163 Synthetic
80 989 81.5 711 11 AY678265 Synthetic
81 987 81.3 711 11 AY678264 Synthetic
82 985 81.1 681 2 AY786536 Discosoma
83 973 80.1 711 11 AY678266 Synthetic
84 972 80.1 681 2 AY786537 Discosoma
85 960 79.1 711 11 AY678267 Synthetic

ALIGNMENTS

RESULT 1
LOCUS AR183915 678 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 6 from patent US 6342379.
ACCESSION AR183915
VERSION AR183915.1 GI:20227884
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 678)
AUTHORS Tseien,R.Y. and Gonzalez,J.E. III.
TITLE Detection of transmembrane potentials by optical methods
JOURNAL Patent: US 6342379-A 6 29-JAN-2002;
FEATURES
    Location/Qualifiers
        source
            1..678
            /organism="Unknown"
            /mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.4e-114 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-006-922A-12 (1-225) x AR183915 (1-678)
QY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20
DB 1 ATGAGGCTCTTCAGAAATGTTATCAAGAGATTCATGAGGTTTAAAGTTTCGATGGAAGGA 60
QY 21 ThrValAenGIyH3sGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB 61 ACGGTCATGGGCAAGCTTTGAATTGAAAGCGCAAGAGAGGGAGCCATTACGAAGGC 120
QY 41 HisAsnThrValLyLeuLyValThrLyGlyGlyProLeuProPheAlaTyrAspIle 60
DB 121 CACAAATACCGTAAAGCTTAAAGTAAAGATTCATGAGGATTTGGGATATT 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLySHsProAlaAspIlePro 80
DB 181 TTGTCAACCAATTTCAAGTATGAAAGCAAGGATATGTCAGACACCTCGCAGACATACCA 240
QY 81 AspTyrLySLyLeuSerPheProGluGlyPheLySTrpgLuarGValMetAsnPheGlu 100
DB 241 GACATATAAAAGCGTGTCTTTCTCGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyPheIleTyr 120
DB 301 GACCGTGGCGTCTTACTGTAACCAAGGATTCAGATTGGCAGATGGCTGTTTCATCTAC 360
QY 121 LySTrValLySLyLeuSerPheProGluGlyPheLySTrpgLuarGValMetAsnPheThr 140
DB 361 AAGGTCAAGTTCATTTGGGATCTTTCTTCGATGAGCCTGTTATGCAAAAGAGGACA 420
QY 141 MetGlyTTrpgLuarLaserThrGluArgLeuTyrProArgAspGlyValIleuLyGlyGlu 160
```

```
DB 421 ATGGCGTGGAGAGCCAGACACTGAGCGTTTGATCTCGTGTATGGCGTGTGAAGAGAG 480
QY 161 IleHisLySalLeuLySLyLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIle 180
DB 481 ATTATATAAGGCTCTGAAAGCTGAAAGAGAGCGTGTCTATTACCTAGTTGAATTCAAAAGTAT 540
QY 181 TyrMetLalLySLySLyProValGlnLeuProGlyTyrTyrTyrValAspSerLyLeuAsp 200
DB 541 TCAATGGCAAGAGAGCCCTGTGACACTACAGGGTACTACTATGTTGACTCCAAACTGGAT 600
QY 201 IleThrSerHisAenGIyAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB 601 ATTAACAAGCCCAACAAGCAAGACTATACATCGTTGAGCAGATGAAAGAACCGAGGAGCAGC 660
QY 221 HisHisLeuPheLeu 225
DB 661 CACCATCTGTTCCCT 675

RESULT 2
LOCUS AX172854 678 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 6 from Patent WO0142211.
ACCESSION AX172854
VERSION AX172854.1 GI:14597903
KEYWORDS
SOURCE Discosoma sp.
ORGANISM Discosoma sp.
REFERENCE 1
AUTHORS Tseien,R.Y. and Gonzalez,J.E.
TITLE Detection of transmembrane potentials by optical methods
JOURNAL Patent: WO 0142211-A 6 14-JUN-2001;
FEATURES
    Location/Qualifiers
        source
            1..678
            /organism="Discosoma sp."
            /mol_type="unassigned DNA"
            /db_xref="taxon:86600"

ORIGIN
Alignment Scores:
Pred. No.: 1.4e-114 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-006-922A-12 (1-225) x AX172854 (1-678)
QY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20
DB 1 ATGAGGCTCTTCAGAAATGTTATCAAGAGATTCATGAGGTTTAAAGTTTCGATGGAAGGA 60
QY 21 ThrValAenGIyH3sGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB 61 ACGGTCATGGGCAAGCTTTGAATTGAAAGCGCAAGAGAGGGAGCCATTACGAAGGC 120
QY 41 HisAsnThrValLyLeuLyValThrLyGlyGlyProLeuProPheAlaTyrAspIle 60
DB 121 CACAAATACCGTAAAGCTTAAAGTAAAGGAGGAGCCTTTGCCATTTGGGATATT 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLySHsProAlaAspIlePro 80
DB 181 TTGTCAACCAATTTCAAGTATGAAAGCAAGGATATATGTCAGACACCTCGCAGACATACCA 240
QY 81 AspTyrLySLyLeuSerPheProGluGlyPheLySTrpgLuarGValMetAsnPheGlu 100
DB 241 GACATATAAAAGCGTGTCTTTCTCGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyPheIleTyr 120
```



Db 301 GACGGTGGCTGTTACTGTAAACCAAGATTCCAGTTTGACAGATGCGTGTTCATCTAC 360  
Qy 121 LysValIysPheIleIleGlyValIAsnPheProSerAspGlyProValIleMetGlnLysLeuThr 140  
Db 361 AAGGTCAAGTTCATTGGCGTGAATCTTCTCCATGAGCCTGTTATACAAAGAGCA 420  
Qy 141 MetGlyTyrGluIAserThrGluArgLeuTyrProArgAspGlyValIleuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCAGACACTGAGCGTTGTATCTCTGTATGCGGTGTGAAGAGAG 480  
Qy 161 ILeHsIlyValIleuLysLeuLysAspGlyGlyIAsnTyrLeuValIleuPheLysSerIle 180  
Db 481 ATTCATTAAGGCTCTGAACCTGAAGACGAGTGTCTATTAAGTGAATTCAAAAGTAT 540  
Qy 181 TyrMetAlaLysPheProValIleuPheProGlyTyrTyrTyrValIAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAACCTGTGACGTTACAGGCTACTATGTGTGACTCCAAACTGAT 600  
Qy 201 ILeThrSerHsIAsnGluAspTyrThrIleValIleuGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATTAACAACCCAAAGAGACTATCAATCGTTGAGCAGTATGAAGAAGACGAGGAGCG 660  
Qy 221 HsHsIleuPheLeu 225  
Db 661 CACCATCTGTTCTT 675

RESULT 3  
AX207715 678 bp DNA linear PAT 31-AUG-2001

LOCUS AX207715 Sequence 13 from Patent WO0157242.  
DEFINITION AX207715  
ACCESSION AX207715  
VERSION AX207715.1 GI:15422399  
KEYWORDS

## SOURCE

ORGANISM  
Discoosoma sp.  
Discoosoma sp.  
Bukeriyota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discoomatidae; Discoosoma.

REFERENCE  
1 Stack J.H., Whitney M., Cubitt A.B. and Pollok B.A.  
METHODS OF PROTEIN DESTABILIZATION AND USES THEREOF  
Patent: WO 0157242-A 13 09-AUG-2001;  
JOURNAL  
Aureora Biosciences Corporation (US)

## FEATURES

source  
1..678  
/organism="Discoosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AX207715 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValIArgMetGluGly 20  
Db 1 ATGAGGCTCTCCAAAGATGTTATCAAGAGATTCATAGGTTTAAAGTTTCGCATGGAAGCA 60  
Qy 21 ThrValIAsnGlyHsIleGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTCATGCGGACCAAGATTGAATAAGAAAGCGAAGGAGGAGGCGCATACGAAGGC 120  
Qy 41 HsAsnThrValLysLeuLysValIThrLysGlyGlyProLeuPheAspIle 60  
Db 121 CACAAATCCCTAAAGCTTAAGTAAAGCAAGGCGGAGCCTTTGCAATTTGCTGGATATT 180  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValIThrLysHsIleProAlaAspIlePro 80

Db 181 TTTCACCAACAATTCAGTATGAAGCAAGGTATATGTCAAGCAACCTGCCGACATACCA 240  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValIleAsnPheGlu 100  
Db 241 GACTATAAAGAGCTGATCTTCTGAAGAGATTTAATGGGAAGGTCATGAACCTTGAA 300  
Qy 101 AspGlyGlyValIleThrValIThrGlnAspSerSerLeuGlnAspGlyCyAspHsIleTyr 120  
Db 301 GACGGTGGCTGTTACTGTAAACCAAGATTCCAGTTTGACAGATGCGTGTTCATCTAC 360  
Qy 121 LysValIysPheIleIleGlyValIAsnPheProSerAspGlyProValIleMetGlnLysThr 140  
Db 361 AAGGTCAAGTTCATTGGGTGAACCTTCTCCATGAGCCTGTTATACAAAGAGCA 420  
Qy 141 MetGlyTyrGluIAserThrGluArgLeuTyrProArgAspGlyValIleuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCAGACACTGAGCGTTGTATCTCTGTATGCGGTGTGAAGAGAG 480  
Qy 161 ILeHsIlyValIleuLysLeuLysAspGlyGlyIAsnTyrLeuValIleuPheLysSerIle 180  
Db 481 ATTCATTAAGGCTCTGAACCTGAAGACGAGTGTCTATTAAGTGAATTCAAAAGTAT 540  
Qy 181 TyrMetAlaLysPheProValIleuPheProGlyTyrTyrTyrValIAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAACCTGTGACGTTACAGGCTACTATGTGTGACTCCAAACTGAT 600  
Qy 201 ILeThrSerHsIAsnGluAspTyrThrIleValIleuGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATTAACAACCCAAAGAGACTATCAATCGTTGAGCAGTATGAAGAAGACGAGGAGCG 660  
Qy 221 HsHsIleuPheLeu 225  
Db 661 CACCATCTGTTCTT 675

## RESULT 4

AX233581 678 bp DNA linear PAT 11-SEP-2001

LOCUS AX233581 Sequence 5 from Patent WO0162919.  
DEFINITION AX233581  
ACCESSION AX233581  
VERSION AX233581.1 GI:15593305  
KEYWORDS

## SOURCE

ORGANISM  
Discoosoma sp.  
Discoosoma sp.  
Bukeriyota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discoomatidae; Discoosoma.

REFERENCE  
1 Nelson D., Zamatra E. and Tsien R.  
MODIFIED FLUORESCENT PROTEINS  
Patent: WO 0162919-A 5 30-AUG-2001;  
JOURNAL  
Aureora Biosciences Corporation (US)

## FEATURES

source  
1..678  
/organism="Discoosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"  
/note="red"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AX233581 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValIArgMetGluGly 20  
Db 1 ATGAGGCTCTCCAAAGATGTTATCAAGAGATTCATAGGTTTAAAGTTTCGCATGGAAGCA 60  
Qy 21 ThrValIAsnGlyHsIleGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

|||||  
Db ACGGTCATGGGCAAGATTGAAATAGAGCGAAGAGGGGAGCCCATACGAAGGC 120  
Qy 41 HisbenthrValIysLeuLysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACATATACCGTAAGAGCTTAAGGTAAACCAAGGGGAGACCTTGGCACTTGGGATATT 180  
Qy 61 LeuSerProGlnPheGlnTrpGlySerIysValTyrValIysHisProAlaAspIlePro 80  
Db 181 TTGTCCACCAACATTTCACTATAGAAAGCAAGGTATATGTCCAAAGCACCCTGCCAGCATACCA 240  
Qy 81 AspTyrIysLysLeuSerPheProGlnGlyPheIysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAGAGCTGCTATTCCTGAGAGATTAAATGGAAAGGGCATGAACCTTGAA 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTGGTAACTGTAAACCAAGATTCCAGATTGACAGATGGCTGTTTCATCTAC 360  
Qy 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleTyr 140  
Db 361 AAGGTCAAGTTCATTTGGCTGAACCTTCTCCAGATGACCTGTTATGCAAAAGAGACA 420  
Qy 141 MetGlyTrpGluIAserThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu 160  
Db 421 ATGGGCTGGAGAGCCAGCACTGAGGCTTGTATCTCGTATGGCGTGTGAAAGAGAG 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGlnPheLysSerIle 180  
Db 481 ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCACTTCAAGTGAATTCAAAAGATT 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200  
Db 541 TACATGGGAAGAACCTCTGTCAAGCTACAGGGTACTATGTTGACTCCAACTGGAT 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATACCAAGCCCAACGAAGACTATACATCGTTGAGAGTGAAGAAGACGAAGGAGACGC 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCTT 675

RESULT 5  
AX370404 678 bp DNA linear PAT 16-FEB-2002  
LOCUS Sequence 1 from Patent WO0196373.  
ACCESSION AX370404  
VERSION AX370404.1 GI:18857490  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
other sequences; artificial sequences.  
REFERENCE  
1 Pradkov,A.F. and Terzikikh,A.  
TITLE Fluorescent timer proteins and methods for their use  
JOURNAL Patent: WO 0196373-A 1 20-DEC-2001;  
Clontech Laboratories Inc. (US)  
FEATURES  
SOURCE  
1. .678  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="variant of sequence from Discosoma sp."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 6

US-10-006-922a-12 (1-225) x AX370404 (1-678)  
Qy 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGGGCTCTTCCAAAGACGTATCAAGAGTTCATGCGCTTCAAGAGTGGCATGAGAGGC 60  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyValGluGlyArgProTyrGlyGly 40  
Db 61 ACCGTGAACGGCCACGAGTTGAGATGAGAGGCCAGAGGCCAGGCCCTCTTACAGAGGC 120  
Qy 41 HisbenthrValIysLeuLysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAACACCGGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTCGCTGGAGACATC 180  
Qy 61 LeuSerProGlnPheGlnTrpGlySerIysValTyrValIysHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCAGTTCAGTCCAGGCTCCAGAGGTGATGTGAAGCACCCCGCGACATCCCC 240  
Qy 81 AspTyrIysLysLeuSerPheProGlnGlyPheIysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAAGAGCTGCTCTTCCCGAGGCTTCAAGTGGAGGCCCGTGTATGAACCTTGAG 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGGCGTGGTGAACCGTGAACCAAGACTCTCCCTGCGAGACGGCTGCTCATCTAC 360  
Qy 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleTyr 140  
Db 361 AAGGTGAAGTTCATCGGCGTAACCTTCTCCAGAGGCCCGCGTGAAGACGAAGAAGACC 420  
Qy 141 MetGlyTrpGluIAserThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu 160  
Db 421 ATGGGCTGGAGAGGCTCTCAACGAGGCTGTATCCCGCGCAGCGGTGTGAAGGGGAG 480  
Qy 161 IleHisLysAlaLeuLysLysAspGlyGlyHisTyrLeuValGlnPheLysSerIle 180  
Db 481 ATCCAAAGGCCCTGAAGTGAAGAGCGGCGCCACTACCTGTGGAGTTCANACTCATC 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200  
Db 541 TACATGGCAAGAACCCGTGTCAAGCTCCGGCTACTACTACGTGAGCTCCAACTGGATC 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACTCCCAACGAAGACTACACATCGTGGAGAGTGAAGAGCCGACCGAGGGCGGC 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCACCTGTTCTTG 675

RESULT 6  
AX824725 678 bp DNA linear PAT 11-DEC-2003  
LOCUS Sequence 7 from Patent WO02068459.  
ACCESSION AX824725  
VERSION AX824725.1 GI:39750591  
KEYWORDS  
SOURCE  
ORGANISM  
Discosoma sp.  
Discosoma sp.  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
REFERENCE  
1  
TITLE Non aggregating fluorescent proteins and methods for using the same  
JOURNAL Patent: WO 0206845-A 7 06-SEP-2002;  
FEATURES  
SOURCE  
1. .678  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"

ORIGIN



```

Db      474  ATGGGCTGGGAAGCCAGACGTGAGGCTTTGATCCTCGATGAGGCTGTGTAAGAGAG 533
Qy      161  ILeHsIbYsAlaLeuLybLeuLybAapGlyGlyVhiSTyRLeuValGluPhelySerIle 180
Db      534  ATTCTAAGGCTCTGAGACCTGAAAGACGGTGTCTATTACCTAGTGAATTCMAAGTATT 593
Qy      181  TyMeCaIaLybLybProValGlnLeuProGlyTyTyTyTyValAapSerLybLeuAap 200
Db      594  TACATGGCAAGAAAGCCTGTGCACTACCAAGGTACTATGTTGACTCCAAACTGGAT 653
Qy      201  ILeThSerHIsaAngIuaSPTyThrIleValGluGlnTyrgIuaArgThrgIuLyArg 220
Db      654  ATAAACAAGCCACAAGAGACTATACATCGTTGAGCATGTAAGAAAGAACGAGGAGCG 713
Qy      221  HIsHIsLeuPhelLeu 225
Db      714  CACCATCTGTTCTT 728

RESULT 8
AR636082
LOCUS      AR636082      859 bp      DNA      linear      PAT 14-FEB-2005
DEFINITION Sequence 11 from patent US 6852849.
ACCESSION  AR636082
VERSION     AR636082.1  GI:59795931
KEYWORDS
SOURCE      .
ORGANISM    Unknown.
REFERENCE    1 (bases 1 to 859)
AUTHORS      Tsien, R.Y. and Campbell, R.E.
TITLE         Non-oligomerizing tandem fluorescent proteins
JOURNML      Patent: US 6852849-A 11 08-FEB-2005;
              The Regents of the University of California; Oakland, CA
FEATURES
    source
        1..859
        /organism="unknown"
        /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      1,87e-114      Length:      859
Score:          1214.00      Matches:      225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             Gaps:      0

US-10-006-922A-12 (1-225) x AR636082 (1-859)
Qy      1  MeArGSeSerLybAaNVaIleLybGluPhMeArGPhelybValAArgMeGluGly 20
Db      54  ATGAGGCTTCCAGAAATGTAATCAAGAGATTCATGAGGTTTAAAGTTTCGCAAGAGAA 113
Qy      21  ThrValaAngLybHIsGluPhGluIleGluGlyGluGlyArgProTyrgIuGly 40
Db      114  ACGGTCATGGGCAAGAGTTGAATATAGAGCGCAAGAGAGGCGGCATATAGAAAGC 173
Qy      41  HIsaAnThrValLybLeuLybValThrLybGlyGlyProLeuProPhaIaITpAspIle 60
Db      174  CACATATCCGTAAGACTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGGCTTGGGATATT 233
Qy      61  LeuSerProGlnPhGlnTyrgLySerLybValTyRValLybHIsProAlaAspIlePro 80
Db      234  TTGTACACCAATTTCAATATGGAAGCAAGGTATATGTCAAGCACCTTCGCGACATACCA 293
Qy      81  AspTyRlybLybLeuSerPheProGluGlyPheLybTyRGIuaArgValMeArbPheGlu 100
Db      294  GACATATAAAAGCTCTCATTTCTCTGAAGATTAAATGGAAAGGCTCATGAACCTTTGAA 353
Qy      101  AapGlyGlyValValThrValThrgIuaPseSerSerLeuGlnAapGlyCybPheIleTyR 120
Db      354  GACGGTGGCGTCTGTTACTGTAAACCAAGATTCACGTTGACAGATGCGCTGTTCACTTAC 413

```

```

Qy      121  LybValLybPheIleGlyValIaAsnPheProSerAapGlyProValMeGlnLybLybThr 140
Db      414  AAGGTCAAGTTCAATTTGGCGTGAACCTTTCCATGAGGACTGTTATACAAAGAAAGCA 473
Qy      141  MeGlyTyRGIuaIaSerThrgIuaArgLeuTyRProArgAapGlyValIleuLybGlyIu 160
Db      474  ATGGGCTGGGAAGCCAGACACTGAGCGTTTGATCCTCGTATGGCGTGTGTAAGAGAG 533
Qy      161  ILeHsIbYsAlaLeuLybLeuLybAapGlyGlyVhiSTyRLeuValGluPhelySerIle 180
Db      534  ATTCTAAGGCTCTGAGACCTGAAAGACGGTGTCTATTACCTAGTGAATTCMAAGTATT 593
Qy      181  TyMeCaIaLybLybProValGlnLeuProGlyTyTyTyTyValAapSerLybLeuAap 200
Db      594  TACATGGCAAGAAAGCCTGTGCACTACCAAGGTACTATGTTGACTCCAAACTGGAT 653
Qy      201  ILeThSerHIsaAngIuaSPTyThrIleValGluGlnTyrgIuaArgThrgIuLyArg 220
Db      654  ATAAACAAGCCACAAGAGACTATACATCGTTGAGCATGTAAGAAAGAACGAGGAGCG 713
Qy      221  HIsHIsLeuPhelLeu 225
Db      714  CACCATCTGTTCTT 728

RESULT 9
AR670157
LOCUS      AR670157      859 bp      DNA      linear      PAT 13-JUN-2005
DEFINITION Sequence 11 from patent US 6900304.
ACCESSION  AR670157
VERSION     AR670157.1  GI:67608896
KEYWORDS
SOURCE      .
ORGANISM    Unknown.
REFERENCE    1 (bases 1 to 859)
AUTHORS      Tsien, R.Y., Ting, A.Y. and Zhang, J.
TITLE         Emission ratiometric indicators of phosphorylation
JOURNML      Patent: US 6900304-A 11 31-MAY-2005;
              The Regents of the University of California; Oakland, CA
FEATURES
    source
        1..859
        /organism="unknown"
        /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      1,87e-114      Length:      859
Score:          1214.00      Matches:      225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             Gaps:      0

US-10-006-922A-12 (1-225) x AR670157 (1-859)
Qy      1  MeArGSeSerLybAaNVaIleLybGluPhMeArGPhelybValAArgMeGluGly 20
Db      54  ATGAGGCTTCCAGAAATGTAATCAAGAGATTCATGAGGTTTAAAGTTTCGATGGAAGA 113
Qy      21  ThrValaAngLybHIsGluPhGluIleGluGlyGluGlyArgProTyrgIuGly 40
Db      114  ACGGTCATGGGCAAGAGTTGAATATGAGCGCAAGAGAGGCGGAGGCATATAGAAAGC 173
Qy      41  HIsaAnThrValLybLeuLybValThrLybGlyGlyProLeuProPhaIaITpAspIle 60
Db      174  CACATATCCGTAAGACTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGGCTTGGGATATT 233
Qy      61  LeuSerProGlnPhGlnTyrgLySerLybValTyRValLybHIsProAlaAspIlePro 80
Db      234  TTGTACACCAATTTCAATATGGAAGCAAGGTATATGTCAAGCACCTTCGCGACATACCA 293
Qy      81  AspTyRlybLybLeuSerPheProGluGlyPheLybTyRGIuaArgValMeArbPheGlu 100

```

```

Db      294  |||||GACTATAAAAAGCTGTCATTTCTCGAAGCATTTAAATGGAAAGGTCATGAACCTTTGAA 353
Qy      101  |||||AappgylgylvalThrValThrglnaSpSerSerLeuGlnaSpGlyCyaphe1leTyr 120
Db      354  |||||GACGGTGGCGTCTGACTGTAACCGAGATTCACATTTGCGAGATGGCTTTTCACTCAC 413
Qy      121  |||||LysValIysPhe1leGlyValaAsnPheProSerAaspGlyProValMetGlnIlyIysThr 140
Db      414  |||||AAGGTCAGTTCATTTGGCGTGAATCTTCTTCCTCCATGACCTGTTATCAAAAGAGACA 473
Qy      141  |||||MetGlyTPGluIaSerThrGluArgLeuTyrProArgaSpGlyValIleuIysGlyIu 160
Db      474  |||||ATGGCTGGGAAGCCAGACCTGAAGCGTGTGATCTCTCGATGGGCTGTGAAGAGAG 533
Qy      161  |||||IleHlsValaLeuIysLeuIysaSpGlyIyHlsTyrLeuValGlnPheIysSerIle 180
Db      534  |||||ATTCTAAAGCTCTGAAGCTGMAAGACGGTGTCTTACTTACCTGAAATTCAAAAGTATT 593
Qy      181  |||||TyrMetAlaIysLysProValGlnLeuProGlyTyrTyrTyrValaAspSerIysLeuAsp 200
Db      594  |||||TACATGGCAAGAACCTGTGACGCTACAGGCTACTATGTTGACTCCAAACTGAT 653
Qy      201  |||||IleThrSerHlsAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGlnIyArg 220
Db      654  |||||ATTAACAGCCACAAAGACATTAACAATCGTTGAGCAGTATGAAGAAGACGAGGACGC 713
Qy      221  |||||HlsHlsLeuPheLeu 225
Db      714  |||||CACCATCTGTTCTT 728

```

```

RESULT 10
LOCUS      AX463698      859 bp      DNA      linear      PAT 15-JUL-2002
DEFINITION Sequence 12 from Patent WO0248338.
ACCESSION  AX463698
VERSION     AX463698.1 GI:21886457
KEYWORDS
SOURCE
ORGANISM   Discosoma sp.
            Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;
            Corallimorpharia; Discosomatidae; Discosoma.

```

```

REFERENCE 1
AUTHORS   Lichtenberg-Frac, H.
TITLE      Yeast strain for testing the geno- and cytotoxicity of complex
            environmental contamination
JOURNAL    Patent: WO 0248338-A 12 20-JUN-2002;
            Lichtenberg-Frac, Heila (DE)
FEATURES   source
            1..859
                /organism="Discosoma sp."
                /mol_type="unassigned DNA"
                /db_xref="taxon:8660"
                54..731
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD42147.1"
                /db_xref="GI:21886458"
                /translation="MRSSKNVTKERPRFKYRMGEYVNGHFEIIEGREGRPYEGNHTV
                KIKVTKGGLPFPAMDILSPQFGSGKYVYVHHPADIDYKLSFPBEGFKERVMANPEDG
                GVATVQDSLSLDGCPYIKVKEIGVNPBSPDMQKTKGMEASERLTPROGVKGE
                IHKALTKDGCHYLVEFKSIYNAKKVPVQPGYYVYDSKLDITSHNEDYTIIVQYERTE
                GRHHLFL"

```

```

ORIGIN
Alignment Scores:
Pred. No.:      1..87e-114      Length:      859
Score:          1214.00         Matches:      225
Percent Simlarity: 100.00%      Mismatches:  0
Best Local Simlarity: 100.00%      Indels:      0
Query Match:    100.00%         Gaps:        0
DB:

```

```

US-10-006-922a-12 (1-225) x AX463698 (1-859)
Qy      1  |||||MetArgSerSerIysaAsnValIleIysGlnPheMetArgPheIysValArgMetGlnGly 20
Db      54  |||||ATGAGGCTCTCCAAAGATGTTATCAAGAGATTCAAGGTTAAAGTTGCGATGGAAGGA 113
Qy      21  |||||ThrValaAsnIyHlsGluPheGluIleGlnIyGlnIyGlnIyArgProTyrGlnIy 40
Db      114  |||||ACGGTCAATGGGCAAGACTTTGAATTAAGACCGAAGAGAGGGAGGCCATTACGAAGC 173
Qy      41  |||||HlsAsnThrValIysLeuIysValaThrIysGlyIyProLeuProPheAlaTyrAspIle 60
Db      174  |||||CACAAATACGTAAGAGCTTAAGATTAACCAAGGGGAGCTTTGGCATTTGCTGGGATATT 233
Qy      61  |||||LeuSerProGlnPheGlnTyrGlySerIysValIyValIysPheIysProAlaAspIlePro 80
Db      224  |||||TTGTACACCAATTCATTAAGTATGAAGCAAGGATATATGCAAGCACCTGCGACATACCA 293
Qy      81  |||||AapTyrIysLysLeuSerPheProGlnIyPheIysTPGluArgValaMetAsnPheGlu 100
Db      294  |||||GACTATAAAAAGCTGTCATTTCTCGAAGCATTTAAATGGAAAGGTCATGAACCTTTGAA 353
Qy      101  |||||AappgylgylvalThrValThrglnaSpSerSerLeuGlnaSpGlyCyaphe1leTyr 120
Db      354  |||||GACGGTGGCGTCTGACTGTAACCGAGATTCACATTTGCGAGATGGCTTTTCACTCAC 413
Qy      121  |||||LysValIysPhe1leGlyValaAsnPheProSerAaspGlyProValMetGlnIlyIysThr 140
Db      414  |||||AAGTCAAGTTCATTTGGCGTGAATCTTCTTCCTCCATGACCTGTTATCAAAAGAGACA 473
Qy      141  |||||MetGlyTPGluIaSerThrGluArgLeuTyrProArgaSpGlyValIleuIysGlyIu 160
Db      474  |||||ATGGCTGGGAAGCCAGACCTGAAGCGTGTGATCTCTCGATGGGCTGTGAAGAGAG 533
Qy      161  |||||IleHlsValaLeuIysLeuIysaSpGlyIyHlsTyrLeuValGlnPheIysSerIle 180
Db      534  |||||ATTCTAAAGCTCTGAAGCTGMAAGACGGTGTCTTACTTACCTGAAATTCAAAAGTATT 593
Qy      181  |||||TyrMetAlaIysLysProValGlnLeuProGlyTyrTyrTyrValaAspSerIysLeuAsp 200
Db      594  |||||TACATGGCAAGAACCTGTGACGCTACAGGCTACTATGTTGACTCCAAACTGAT 653
Qy      201  |||||IleThrSerHlsAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGlnIyArg 220
Db      654  |||||ATTAACAGCCACAAAGACATTAACAATCGTTGAGCAGTATGAAGAAGACGAGGACGC 713
Qy      221  |||||HlsHlsLeuPheLeu 225
Db      714  |||||CACCATCTGTTCTT 728

```

```

RESULT 11
LOCUS      AX233584      681 bp      DNA      linear      PAT 11-SEP-2001
DEFINITION Sequence 8 from Patent WO0162919.
ACCESSION  AX233584
VERSION     AX233584.1 GI:15593307
KEYWORDS
SOURCE
ORGANISM   synthetic construct
            synthetic construct
            other sequences; artificial sequences.
REFERENCE 1
AUTHORS   Nelson, D., Zamaira, E. and Tsien, R.
TITLE      Modified fluorescent proteins
JOURNAL    Patent: WO 0162919-A 8 30-AUG-2001;
            Aurora Biosciences Corporation (US)
FEATURES   source
            1..681
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Mutant Anthozoan red fluorescent protein"
                1..681
                /note="unnamed protein product"

```

```
/codon_start=1
/translation_table=11
/protein_id="CAC69734.1"
/db_xref="GI:1593308"
/translation="MVRSSKNVKEFMKFKVMEGTVNGHFEIEGSEGRPYEGHNT
VKLKVTKGSLPFPAMDLSKPOQYGSKYVVGHPDIDPYKYSPEEGKMRVNFED
GGVTVNQDSSLQDGCPIYKVKFIQVNPSPSGPQWOKTQMWBSTERYLRDQVLKG
EIKHAKLKQSGHLYVERKSIYMAKKVPQLPGYYVDSKLDITSHNDYITVEQYERT
EGRHHLPL"
```

## ORIGIN

## Alignment Scores:

Pred. No.:	3,61e-114	Length:	681
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922a-12 (1-225) x AX233584 (1-681)

```
QY      1 MetArgSerSerLySAenValIleLyGluPheMetArgPheLyValArgMetGluGly 20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      4 GTGAGGAGCAGCAAGAACGTGATCAAGAGTTCAATGAGTTCAAGGTCCGATGAGGCG 63

QY      21 ThrValaEngLyHISgluPheGluIleGluGlyGluGlyArgProTyrgLugly 40
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      64 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGCAGGCGCAGGCGCCTTACGAGGCG 123

QY      41 HisAsnThrValLySLeuLySValThrLySGLyProLeuProPheAlaTrpAspIle 60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      124 CACAACACCGTGAAGCTTAAAGTACCAAGGCGGCGCCCTTGCCTTGGGACATC 183

QY      61 LeuSerProGlnPheGlnTyrgLySerLySValTyrgValySHISProAlaAspIlePro 80
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      184 CTGAGCCCCCAGTTCCAGTACGCGCAGCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 243

QY      81 AspTyrlLySLeuSerPheProGluGlyPheLySTrpGlyuArgValMetAsnPheGlu 100
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      244 GACTACAAAGAAAGCTGAGCTTCCCGAGGCTTCAAGTGGAGAGGTATGAACCTTCGAG 303

QY      101 AspGlyGlyValIleThrValIleThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyrg 120
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      304 GACGCGCGCGTGGTGAACCGTGAACCAAGACAGCAAGCTGCGAGCGCTTCACTAC 363

QY      121 LySValLySPhelIleGlyValIleAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      364 AAGGTGAAGTTCACTCGCGGTGAACCTTCCCGACGCGCGCCCGTGAATGCAAGAAAGAC 423

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrgProArgAspGlyValLeuLySGLyGlu 160
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      424 ATGGGCTGGAGGCTCCACCGAGGCGCTGTACCCCGCAGCGGCGGTGTAAAGGCGAG 483

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHISLyTrLeuValIlePheLySerIle 180
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      484 ATCCAAAGGCGCTGAAGCTGAAGGACGCGCGCACTACCTGGTGAAGTTCAAGTCCATC 543

QY      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrgTyrgValIleAspSerLySLeuAsp 200
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      544 TACATGGCCAAAGAAAGCCGCTGACGCTGCCGCTACTACTAGTGAAGTCCAAAGCTGAGC 603

QY      201 IleHisSerHISaEngLySAspTyrgThrIleValGluGlnTyrgIuArgThGluGlyArg 220
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      604 ATCCACGAGCCCAACGAGAGCTACACATCGTGAAGCATGTACGAGAGGACCGAGGCGAG 663

QY      221 HisHisLeuPheLeu 225
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      664 CACCACTGTTCTG 678

RESULT 12
AX233627 713 bp DNA linear PAT 11-SEP-2001
LOCUS AX233627
DEFINITION Sequence 51 from Patent WO0162919.
```

```
ACCESSION AX233627
VERSION AX233627.1 GI:15933330
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 Nelson,D., Zamaiera,E. and Tsien,R.
TITLE Modified fluorescent proteins
JOURNAL Patent: WO 0162919-A 51 30-AUG-2001;
Autora Biosciences Corporation (US)
location/Qualifiers
FEATURES
source
```

## FEATURES

## source

```
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Mutant Anthozoan red fluorescent protein"
```

## ORIGIN

## Alignment Scores:

Pred. No.:	3,81e-114	Length:	713
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922a-12 (1-225) x AX233627 (1-713)

```
QY      1 MetArgSerSerLySAenValIleLyGluPheMetArgPheLyValArgMetGluGly 20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      22 GTGAGGAGCAGCAAGAACGTGATCAAGAGTTCAATGAGTTCAAGGTCCGATGAGGCG 81

QY      21 ThrValaEngLyHISgluPheGluIleGluGlyGluGlyArgProTyrgLugly 40
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      82 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGCAGGCGCAGGCGCCTTACGAGGCG 141

QY      41 HisAsnThrValLySLeuLySValThrLySGLyProLeuProPheAlaTrpAspIle 60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      142 CACAACACCGTGAAGCTTAAAGTACCAAGGCGGCGCCCTTGCCTTGGGACATC 201

QY      61 LeuSerProGlnPheGlnTyrgLySerLySValTyrgValySHISProAlaAspIlePro 80
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      202 CTGAGCCCCCAGTTCCAGTACGCGCAGCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 261

QY      81 AspTyrlLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      262 GACTACAAAGAAAGCTGAGCTTCCCGAGGCTTCAAGTGGAGAGGTGTGAACCTTCGAG 321

QY      101 AspGlyGlyValIleThrValIleThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyrg 120
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      322 GACGCGCGCGTGGTGAACCGTGAACCAAGACAGCAAGCTGCGAGCGCTTCACTAC 381

QY      121 LySValLySPhelIleGlyValIleAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      382 AAGGTGAAGTTCACTCGCGGTGAACCTTCCCGACGAGCGGCCCGTGAATGCAAGAAAGAC 441

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrgProArgAspGlyValLeuLySGLyGlu 160
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      442 ATGGGCTGGAGGCTCCACCGAGGCGCTGTACCCCGCAGCGGCGGTGTGAAGGCGAG 501

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHISLyTrLeuValIlePheLySerIle 180
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      502 ATCCAAAGGCGCTGAAGCTGAAGGACGCGCGCACTACCTGGTGAAGTTCAAGTCCATC 561

QY      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrgTyrgValIleAspSerLySLeuAsp 200
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      562 TACATGGCCAAAGAAAGCCGCTGACGCTGCCGCTACTACTAGTGAAGTCCAAAGCTGAGC 621

QY      201 IleHisSerHISaEngLySAspTyrgThrIleValGluGlnTyrgIuArgThGluGlyArg 220
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      622 ATCCACGAGCCCAACGAGAGCTACACCATCGTGAAGCATGTACGAGAGGACCGAGGCGAG 681
```

QY	221	HisHsleuPheleu	225
Db	682	CACCACTGTTCTG	696
RESULT 13			
AR527331			
LOCUS	AR527331	723 bp	DNA
DEFINITION	Sequence 1 from patent US 6723537.		linear
ACCESSION	AR527331		
VERSION	AR527331.1	GI:53914309	
KEYWORDS			
SOURCE	.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 723)		
TITLE	Peelle, B.		
JOURNAL	Directed evolution of protein in mammalian cells		
	Patent: US 6723537-A 1 20-APR-2004;		
	Rigel Pharmaceuticals, Incorporated and Becton, Dickinson and		
	Company: South San Francisco, CA		
FEATURES			
source	1..723		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,886-114	Length:	723
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x AR527331 (1-723)			
QY	1	MeAtgSerSeLyAhAnValIleLySgLuPheMeChgPheLyValArgMeGluGly	20
Db	4	GTGCCCTCTCCAGAACGTCAATCAAGAGTTCATGCGCTTCAAGATGGCAATGAGAGGC	63
QY	21	ThrValaAnGlyVhIeGluPheGluIleGluGlyGluGlyVArpProTyrgluGly	40
Db	64	ACCGGAGAGGCAACAGTTCAAGATCGAGGCGAGAGGCGCGCCCTTACAGAGGC	123
QY	41	HisAenThrValIeLeuLySValThryLySgLyGlyProLeuProPheAlaTrpAspIle	60
Db	124	CACACACCGCTGAGCTGAGGTGACCAAGGCGGCGCCCTCGCTTGGCTGGAGATC	183
QY	61	LeuSerProGluPheGluTyrgLySerLySValTyrrValIyHisProAlaAspIlePro	80
Db	184	CTGTCCCGCCAGTTCAGATGAGGCTCAAGAGTGAAGTGAAGACACCCCGCACATCCCC	243
QY	81	AapTyrtLyLyLeuSerPheProGluGlyPheLySValMetAAnPheGlu	100
Db	244	GACTCAACAAACCTGCTTCCCGAGGCTTCAAGTGAAGCGCTGAATTCGAG	303
QY	101	AapGlyGlyVAlValThrValThrgInAepSerSerLeuGluAAspGlyCySpheIleTy	120
Db	304	GACGCGGCGGTGTGACCTGAGCCAGACGACTCTCCCTGCAAGAGCGGCTTCACTAC	363
QY	121	LyValaLyPheIleGlyValaAnPheProSerAepGlyProValMeGluLySValThr	140
Db	364	AAGGTGAATTCATCGCGCTGAATCTCCCTCCGACGCGCCGCTAATGCAAGAAAGACC	423
QY	141	MeGlyTTrGluAlaSerThrgIuSgLeuTyrrProArgAAspGlyValaLeuLySgLyGlu	160
Db	424	ATGGGCTGGAGGCTTCACCGAGCGCTGTATCCCCGACACGCGTCTGTAAAGGCGAG	483
QY	161	IleHleLyAlaLeuLySValAAspGlyGlyVhIeTyrrLeuValGluPheLySerIle	180
Db	484	ATCCACAGGCGCTTGAAGCTGAAGAGACGCGGCGCACTACTGGTGAAGTTCAAGATATC	543
QY	181	TyMeAlaLyValLySProValGluLeuProGlyTyrrTyrrValaAspSerLySLeuAsp	200

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	ORIGIN
Db	544	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACCTCCAGCTGGAC										
Qy	201	11EhrserHisAnGluAepTyTrHileValGluGlnTyGluAArgThrGluGlyArg										
Db	604	ATCACCTCCCAACAAGAGACTACACCATGTGTGAGAGATGACGAGCGACCGAGGGCCGC										
Qy	221	HisHisLeuPheLeu 225										
Db	664	CACCACCTGTTCTGT 678										
RESULT 14												
Q0882115												
LOCUS	C0882115	Sequence 1 from Patent WO2004083445.										
DEFINITION	C0882115											
ACCESSION	C0882115	GI:54034825										
VERSION	C0882115.1											
KEYWORDS												
ORGANISM												
REFERENCE	1											
AUTHORS	Chavancy, G., Couble, P., Durand, B., Grenier, A.M., Horard, B., Julien, E., Mauchamp, B., Nony, P., Prudhomme, J.C. and Royer, C.											
TITLE	Nucleic acid controlling the expression of a useful polypeptide in the posterio silk glands of a Lepidoptera and application thereof											
JOURNAL	Patent: WO 2004083445-A 1 30-SEP-2004; Centre National de la Recherche Scientifique-CNRS (FR); UNIVERSITE CLAUDE BERNARD - LYON 1 (FR); Institut National de la Recherche Agronomique (INRA) (FR)											
FEATURES	Location/Qualifiers											
source	1..2721											
	/organism="synthetic construct"											
	/mol_type="unassigned DNA"											
	/db_xref="taxon:32630"											
	/note="description de la s quence artificielle : s quence de fusion"											
ALIGNMENT SCORES:												
Pred. No.:	1,95e-113	Length:	2721									
Score:	1210.00	Matches:	224									
Percent Similarity:	100.00%	Conservative:	1									
Best Local Similarity:	99.56%	Mismatches:	0									
Query Match:	99.67%	Indels:	0									
DB:	6	Gaps:	0									
US-10-006-922A-12 (1-225) x CQ882115 (1-2721)												
Qy	1	MetAArgSerSerLyBAenValIleLeuGluPheMetArgPheLyValArgMetGluGly										
Db	2044	GGGCCCTCTCCCAAGAAAGCATCAAGAGATTCAATGCGCTTCAAGAGTGGCATGAGGCG										
Qy	21	ThrValaAnGlyHHisGluPheGluHileGluGlyGluGlyGlyArgProTyGluGly										
Db	2104	ACCGTGAACGGCCACGATTCGAGATTCGAGGCGCAGGGCCGCGCCCTTACGAGGCG										
Qy	41	HisAenThrValLyLeuLeuValThrLyGlyGlyProLeuProPheAlaTrpAseIle										
Db	2164	CACAAACACCGTAAAGCTGAGGTCACAAAGGCGCGCCCTTGCGCTGGGACATTC										
Qy	61	LeuSerProGlnPheGlnTyGlySerTyValTyTrValLyHisAProAlaAspIlePro										
Db	2224	CTGTCCCCCACTTCACTAGCGCTCCCAAGGATGACGTGAAGCACCCCGCGACATCCC										
Qy	81	AspTyTrLyLyLeuSerPheProGluGlyPheLySTPGLuArgValMetAsnPheGlu										
Db	2284	GACTACAAAGAGCTCTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTATGAATTGAG										
Qy	101	AspGlyGlyValValThrValThr										

```
Qy 121 LysValIysPheIleGlyValIasnPheProSerAspGlyProValIleMetGlnLysLeuThr 140
Db 2404 AAGGTGAAGTTCATCGGCGTAACCTCCCTCCGAGGCGCCGTAATGCAAGAGAGACC 2463
Qy 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleLysGlyGlu 160
Db 2464 ATGGGCTGGAGGGCTCCACCGAGGCGCTGTATCCCGCGAGCGGGGTGTCTAGAGGGGAG 2523
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db 2524 ATCCACAAGGCGCTGAACTGAAGAGACGCGCGCACTACCTGTGTGAAGTTCAGATCCATC 2583
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAlaAspSerLysLeuAsp 200
Db 2584 TACATGGCCAGAAAGCCGCTGACGCTGCCGCGCTACTACTACGTGGACTCCAGAGCTGGAC 2643
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 2644 ATCACTCTCCACAAGAGAGATACACCATCTGTGAGAGCATACAGCGCACCGAGGGCGCG 2703
Qy 221 HisHisLeuPheLeu 225
Db 2704 CACCACCTGTTCCTG 2718

RESULT 15
CO882117 CO882117 2772 bp DNA linear PAT 11-OCT-2004
DEFINITION Sequence 3 from Patent WO2004083445.
ACCESSION CO882117
VERSION CO882117.1 GI:54034827
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chavancy, G., Conble, P., Durand, B., Grenier, A.M., Horard, B.,
TITLE Nucleic acid controlling the expression of a useful polypeptide in
JOURNAL the posterior silk glands of a lepidoptera and application thereof
PATENT: WO 2004083445-A 3 30-SEP-2004;
CENTRE NATIONAL De la Recherche Scientifique-CNRS (FR) ; UNIVERSITE
CLAUDE BERNARD - LYON 1 (FR) ; Institut National de la Recherche
Agronomique (INRA) (FR)
FEATURES
SOURCE location/Qualifiers
1..2772
/oranism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description de la s quence artificielle : s quence
de fusion"
ORIGIN
Alignment Scores:
Pred. No.: 2e-113 Length: 2772
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
Gaps: 0
DB: 6
US-10-006-922A-12 (1-225) x CO882117 (1-2772)
```

```
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro 80
Db 2275 CTGTCCCGCCAGTTCAGTACGAGGCTCCAGAGGTGACGTAAGAGACCCCGGACATCCCC 2334
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTTPGluArgValIleMetAsnPheGlu 100
Db 2235 GACTACAAAGAAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGATCACTTGAG 2394
Qy 101 AspGlyGlyValValIleThrValIleGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 2295 GACGCGCGCGTGGACCGTGAACCCAGAGACTCTCCCTGCGAGGAGCGGCTGCTTCACTAC 2454
Qy 121 LysValIysPheIleGlyValIasnPheProSerAspGlyProValIleMetGlnLysLeuThr 140
Db 2455 AAGGTGAAGTTCATCGGCGTAACCTCCCTCCAGAGGCGCCGTAATGCAAGAGAGACC 2514
Qy 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleLysGlyGlu 160
Db 2515 ATGGGCTGGAGGGCTCCACCGAGGCGCTGTATCCCGCGAGCGCGGTGTGAAGGGGAG 2574
Qy 161 IleHisLysAlaLeuLysLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db 2575 ATCCACAAGGCGCTGAACTGAAGAGACGCGCGCACTACCTGTGTGAAGTTCAGATCCATC 2634
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAlaAspSerLysLeuAsp 200
Db 2635 TACATGGCCAGAAAGCCGCTGACGCTGCCGCGCTACTACTACGTGGACTCCAGAGCTGGAC 2694
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 2695 ATCACTCTCCACAAGAGAGATACACCATCTGTGAGAGCATACAGCGCACCGAGGGCGCG 2754
Qy 221 HisHisLeuPheLeu 225
Db 2755 CACCACCTGTTCCTG 2769

RESULT 16
AX463702 AX463702 4692 bp DNA linear PAT 15-JUL-2002
DEFINITION Sequence 16 from Patent WO0248338.
ACCESSION AX463702
VERSION AX463702.1 GI:21886461
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Lichtenberg-Frat, H.
TITLE Yeast strain for testing the geno- and cytotoxicity of complex
JOURNAL environmental contamination
PATENT: WO 0248338-A 16 20-JUN-2002;
LICHTENBERG-FRAT, Hella (DE)
FEATURES
SOURCE location/Qualifiers
1..4692
/oranism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Vektor pDBRed1-N1"
ORIGIN
Alignment Scores:
Pred. No.: 3.79e-113 Length: 4692
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
Gaps: 0
DB: 6
US-10-006-922A-12 (1-225) x AX463702 (1-4692)
```

```
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValIleArgMetGluGly 20
Db 2215 CACAACACCGTGAAGCTGAAGGTGACAAAGGGCGGCGCCCTTGCCCTTGCGGAGCATC 2274
```



```
Db 682 GTGGCTCTCCAGAAAGTCATCAAGAGTTTCATGCGCTTCAGAGTGCATGAGGGC 741
Qy 21 ThrValaAngIyHsAgIuPheGluIleGluGluGluGluGluYArgProTyrgIuGly 40
Db 742 ACCGTGAAGCGCCACAGATTGAGATCGAGGGCGAGGGCGAGGGCGCCCTTACGAGGGC 801
Qy 41 HsAenThrValIyLysLeuLysValIThrIyGlyGlyProLeuProPheAlaSTrAspIle 60
Db 802 CACAACACCGTAGAGCTGAGAGGTGACCAAGGGCGGGCCCTTGCCTTGCCTGGAGCATC 861
Qy 61 LeuSerProGlnPheGlnTyrgIySerIyValIyTrValIyLysProAlaAspIlePro 80
Db 862 CTGTCCCGCCAGTTCCAGTACGCGCTCCAGAGGTGACGTGAAGCACCCGCGACATCCCC 921
Qy 81 AepTyIyLysLysSerPheProGluGlyPheIySTrGluArgValMetAenPheGlu 100
Db 922 GACTACAGAAAGCTCTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCAG 981
Qy 101 AepGlyGlyValIyValIThrValIThrGlnApsSerSerLeuGlnApsGlyCyAphelIeTy 120
Db 982 GACGCGCGCGTGTGATCCGTGACCCAGGACTCTCTCCCTGACAGACGCGCTTCACTTAC 1041
Qy 121 LysValIyPheIleGlyValAenPheProSerAepGlyProValMetGlnIyLysPhe 140
Db 1042 AAGGTGAAGTTCATCGCGCGTGAACCTTCCCTCCAGCGGCCCGTAAATGCAAGAAAGACC 1101
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyProArgApsGlyValIleuysGlyGlu 160
Db 1102 ATGGGCTGGAGGCTCCACCGAGCGCTGTGACCCCGGACGCGGTGTGAAGGGCGAG 1161
Qy 161 IleHsIyAlaLeuLysLeuLysApsGlyGlyHsTyLeuValGluPheLysSerIle 180
Db 1162 ATCCCAAGGCGCTGAGCTGAAGGACGCGCGGCACTACTGTGAGTTCAGATCCATC 1221
Qy 181 TyrMetAlaLysLysProValGluLeuProGlyTyTyTyTyValApsSerIyLysAps 200
Db 1222 TACATGGCCCAAGAGCCCGTGCAGCGCTGCGGCTACTACTGAGTTCAGCGTGCAC 1281
Qy 201 IleThrSerHsAenGluApsTyTyTyTyIleValGluGlnTyrgIuArgThrGluGly 220
Db 1282 ATCACTCTCCACAGAGGACTACACATCTGTGAGAGTACGAGGCGCACGAGGGCCCC 1341
Qy 221 HsHsIleuPheLeu 225
Db 1342 CACCACTCTTCTCTG 1356
RESULT 17
LOCUS AX823860 6893 bp DNA linear PAT 11-DEC-2003
DEFINITION Sequence 8 from Patent WO03070931.
ACCESSION AX823860
VERSION AX823860.1 GI:39750176
KEYWORDS
SOURCE
ORGANISM
SYNTHETIC CONSTRUCT
SYNTHETIC CONSTRUCT
OTHER SEQUENCES; artificial sequences.
REFERENCE
1 Baum, C., Wall, E., Osterberg, W., Klump, H. and Schiedlmeier, B.
AUTHORS Methode for conducting site-specific dna recombination
TITLE Patent: WO 03070931-A 8 28-AUG-2003;
JOURNAL Vision 7 GmbH (DE)
LOCATION/Qualifiers
FEATURES
source
1..6893
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/feature="Beschreibung der kuenstlichen Sequenz: provirale
Plasmid-DNA; retrovirales Cre-Reporterkonstrukt #SF#"
```

```
misc_feature
1..435
/note="Plasmid-Rueckgrat (pUC)"
misc_feature
436..996
/note="MPSV 5'-LTR (Delta1-31)"
primer_bind
997..1014
/note="PBS (primer binding site)"
1015..1560
misc_feature
1561..1599
/note="#loxP1#-Sequenz"
misc_feature
1600..2289
/note="vRed1-Gen aus Discosoma sp."
misc_feature
2290..2343
/note="#loxP2#-Sequenz"
misc_feature
2347..3150
/note="eGFP (tag) -Gen aus Aequorea victoria"
misc_feature
3164..4019
/note="#wPRE#"
3'UTR
4020..4080
misc_feature
4081..4641
/note="3'-LTR"
4642..6893
misc_feature
/note="Plasmid-Rueckgrat (pUC)"
ORIGIN
Alignment Scores:
Pred. No.: 6,06e-113 Length: 6893
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 6 Gaps: 0
US-10-006-922A-12 (1-225) x AX823860 (1-6893)
Qy 1 MetApsSerSerIyAenValIleIyGluPheMetArgPheLysValArgMetGluGly 20
Db 1603 GTGGCTCTCCAGAAAGTCATCAAGAGTTTCATGCGCTTCAAGTGCATGAGGGC 1662
Qy 21 ThrValaAngIyHsAgIuPheGluIleGluGluGluGluYArgProTyrgIuGly 40
Db 1663 ACCGTGAAGCGCCACAGATTGAGATCGAGGGCGAGGGCGAGGGCGCCCTTACGAGGGC 1722
Qy 41 HsAenThrValIyLysLeuLysValIThrIyGlyGlyProLeuProPheAlaSTrAspIle 60
Db 1723 CACAACACCGTAGAGCTGAGAGGTGACCAAGGGCGGCCCTTGCCTTGGAGCATC 1782
Qy 61 LeuSerProGlnPheGlnTyrgIySerIyValIyTrValIyLysProAlaAspIlePro 80
Db 1783 CTGTCCCGCCAGTTCCAGTACGCGCTTCCAGAGTGTACGTGAAGCACCCGCGCATCCCC 1842
Qy 81 AepTyIyLysLysSerPheProGluGlyPheIySTrGluArgValMetAenPheGlu 100
Db 1843 GACTACAGAAAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG 1902
Qy 101 AepGlyGlyValIyValIThrValIThrGlnApsSerSerLeuGlnApsGlyCyAphelIeTy 120
Db 1903 GACGCGCGCGTGTGACCGTGAACCAAGGACTCTCTCCCTGACAGAGCGCGTTCATCTAC 1962
Qy 121 LysValIyPheIleGlyValAenPheProSerAepGlyProValMetGlnIyLysPhe 140
Db 1963 AAGGTGAAGTTCATGCGCGTGAACCTTCCCTCCAGCGGCCCTTGAATGAGAAAGAGCC 2022
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyProArgApsGlyValIleuysGlyGlu 160
Db 2023 ATGGGCTGGAGAGCGCTCCACCGAGCGCTGTACCCCGGAGCGCGGTGTGAAGGGCGAG 2082
Qy 161 IleHsIyAlaLeuLysLeuLysApsGlyGlyHsTyLeuValGluPheLysSerIle 180
Db 2083 ATCCCAAGGCGCTGAGCTGAAGGACGCGCGGCACTACTGTGTGAATTCAGATCCATC 2142
Qy 181 TyrMetAlaLysLysProValGluLeuProGlyTyTyTyTyValApsSerIyLysAps 200
Db 2143 TACATGGCCCAAGAGCCCGTGCAGCGCTGCGGCTACTACTGAGTTCAGATCCAGCTGCAC 2202
Qy 201 IleThrSerHsAenGluApsTyTyTyTyIleValGluGlnTyrgIuArgThrGluGly 220
Db 2203 ATCACTCTCCACAGAGGACTACACATCTGTGAGAGTACGAGGCGCACCGAGGGCCCC 2262
```

	misc_feature	rep_origin	CDS	ORIGIN
	/trnasl_table=11			
	/product="LactI"			
	/protein_id="AA578498.1"			
	/cd_xref="GI:43861236"			
	/trnasl_atlon="MKRPVTLVDYAEVAGVSYQTVSRVVAQASHVSAKTRKREVEANAE			
	LNTIPRVAQQLAGKQSLILGVATSSLLALAPQIVAAIKSRADQGLAVSVSVRS			
	GVACRAAVINELAAQVSGILINYPLDDQALIVEAACTVPAFLFDVSQTPNSII			
	FSHEDETRGLGVEHLVALHQOIALLAGPQLSSVBARLIGMHVYLRNNOILOPAEREG			
	DRNMSGFOOTQMOLNEGIYPTMLVYANDQMLGAMRATRESGLRGVADISVYCYDDT			
	EDSSCYIPBSTTIKQDFRLGQTSVRLQLQLSQGVAKGNQLPVSIVKRTTLAPNT			
	QTASPAALDSLSLQMLARQVRSLSGQ"			
	6520..6553			
	/note="LoxP LB"			
	6992			
	/note="Col B1"			
	/complement(7751..8611)			
	/codon_start=1			
	/trnasl_table=11			
	/product="Amp"			
	/protein_id="AA578499.1"			
	/cd_xref="GI:43861237"			
	/trnasl_atlon="MSIQHRVALIPFFAFLCPVFAHPETLVKVDADQLGARVGY			
	IEIDLNSGKILSEFRPEERPMWSTKVLICGVLISRIDGQQLGRIRHYSDQLVBE			
	YSVTEKHLTDGTVRELCSAALTMSDNTFANLLTTIGAPKELTAVLHMGDHTVSL			
	DRREPIQNEALIPNDERDTPMVAMLTIRLLTGELLTLASROOILDMEADKVAQPL			
	LRRAALPAGRFIADKSGAGERSGIIIALGLPDKPSRIIVITYTSGSATMDERRQLA			
	ETGASLLIKHW"			
	Alignment Scores:			
	Sred. No.: 8.18e-113	Length: 8811		
	Score: 1210.00	Matches: 224		
	Percent Similarity: 100.00%	Conservative: 1		
	Best Local Similarity: 99.56%	Mismatches: 0		
	Query Match: 99.67%	Indels: 0		
	DB: 11	Gaps: 0		
	US-10-006-922A-12 (1-225) x AY569779 (1-8811)			
QY	1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20			
DB	1384 GTGGGCTCCCTCCCAAGAACGTCATCAAGAGATTCTATCGCTTCAAGGTCGCGCATGAGAGGC 1325			
QY	21 ThrValIAsnGlyIHisGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40			
DB	1324 ACCGTGAACCGCCACGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTCAGAGGGC 1265			
QY	41 HisAsnThrValIysLeuIysValIThrIysGlyGlyProLeuProPheIAsnATPAspIle 60			
DB	1264 CACAAACCCGTGAAGCGTGAAGGTGACCAAGAGGGCGGCCCTCGCCCTTCGCTGGGAATC 1205			
QY	61 LeuSerProGlnPheGlnIryGlySerIysValIryValIyHisIAsProAlaAspIlePro 80			
DB	1204 CTGTCCCCCAAGTTCAGATACGGCTCCAGAGTGTACTGTAAAGCACCCCGCGCATCTCCC 1145			
QY	81 AspTyrIlyValLeuSerPheProGluGlyPheIySTrpGluArgValMetAsnPheGlu 100			
DB	1144 GACTACAAAGAGCTGTCTCTCCCGAGGGCTTCAAGTGGGAGCGCGTGAATGAACTTCGAG 1085			
QY	101 AspGlyGlyValValIThrValIThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120			
DB	1084 GACGGCGGCGCTGTGACCGGTGACCCAGAACATCTCTCCCTGAGAGACGGCTGTCATCTAC 1025			
QY	121 IysValIlyAspHeIleGlyValIAsnPheProSerAspGlyProValMetGlnIlyIysThr 140			
DB	1024 AAGGTGAAGTTCATCGGCGTGAATCTCCCTCCGACGGGCCCTGATATGCAGAGAACACC 965			
QY	141 MetGlyTrpGluIAsnThrGluArgLeuTyrProArgAspGlyValIleuIysGlyGlu 160			
DB	964 ATGGGCGTGGAGGCTTCACCGAGCGGCTGTATCCCGCGGACGGCGGTGAAGGGCGAG 905			
QY	161 IleHisIlyValIleuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180			

Db 904 ATCCAGAGCCCTGAGACTGAGAGGACGGCGCCACTACTGCTGAGACTTCAATCCATC 845  
Qy 181 TyrMetAlaIyLySProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAap 200  
Db 844 TACATGGCCAAAGAGCCCGCTGACAGCTGCCGGCTACTACTAGTGGACTCCAGACTGGAC 785  
Qy 201 IleThrSerHisaenGluAapTyrThrIleValGluGlnTyrGluAyrGThglGluIyAyr 220  
Db 784 ATCACTCCCAACAGAGACTACACATCTGGAGCAGTACGAGCGACCGAGGGCGCG 725  
Qy 221 HisHisLeuPheLeu 225  
Db 724 CACCACTGTTCTTG 710  
RESULT 19  
AK663075  
LOCUS AK663075 9320 bp DNA linear PAT 24-MAR-2003  
DEFINITION Sequence 20 from Patent WO02070740.  
ACCESSION AK663075  
VERSION AK663075.1 GI:29169369  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE other sequences; artificial sequences.  
1  
AUTHORS Wiemuelier, L.  
JOURNAL Test system for determining gene toxicities  
Patent: WO 02070740-A 20 12-SEP-2002;  
Wiemuelier, Lisa (DE)  
FEATURES  
source  
1..9320  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:33630"  
/note="plasmid p5-Puro-CMV-(N'-EGFP)-CMV-Red-(EGFP-EJ)"  
1..1592  
/note="Retroviraler Vektor p5NM"  
1617..2216  
/note="unnamed protein product; Puromycin-Resistenzgen aus  
pRetron (Clontech, Palo Alto, CA, USA)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAD80174.1"  
/db\_xref="GI:29169370"  
/translation="MTBYKPTVRLATRDVPRAVRTLAAPADYPATRTTVPDRHIB  
RTTELDELITGVLDIGKVMADGAAVAWTTESVAGAVFAEIGRMAELGSR  
LAAQOMEGLLAPHRKPERAMFLATVGVSPDHOGGIGSAVVLPGLAABRAGVPAFL  
ETSAFRLNLPFYRLGFTVADVECPDRATWCTTKRPA"  
2267..2848  
/note="CMV-Promotor aus pEGFP-N1 (Clontech, Palo Alto, CA,  
USA)"  
2806..3349  
/note="unnamed protein product; N'-EGFP, abgeleitet von  
EGFP aus pEGFP-N1 (Clontech, Palo Alto, CA, USA)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAD80175.1"  
/db\_xref="GI:29169371"  
/translation="WYSKGEELFTGVPLIVELDDGVNHKFSVSGEGSDATYGLT  
LFLCTTGKLPVWPVTLVTLTVGVCFSPYDPHMKQHPFSGAMPEGVYQERTTIFK  
DDGNVYTRAEVRFEGDTLVNRIELKSIDKEQDNILGHLELYNN"  
3374..3392  
/note="Retroviraler Vektor p5NM"  
3411..3992  
/note="CMV-Promotor aus pRed1-N1 (Clontech, Palo Alto,  
CA, USA)"  
4038..4718  
/note="unnamed protein product; Red aus pRed1-N1  
(Clontech, Palo Alto, CA, USA)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAD80176.1"  
/db\_xref="GI:29169372"

misc\_feature  
/translation="WVRSKNVLEFMRFKVRMEGTVNGHEFIEGEGRPREGHNT  
VKIKVTRGSLPFAMDILISPOFGSVYVHPADIIDYKLSFEPGKRMVNFED  
GGVTTVODSLODGCETIKYKIFGNFSPGPMOKKTGMWASTRIYRDLVKG  
EIRHALKIKOGHLYVEFKSTYMAKKEVQLPGLTYVDSKIDTISHNEDYTVQERT  
EGRHLEPL"  
4766..5509  
/note="EGFP-EJ, abgeleitet von EGFP aus pEGFP-N1  
(Clontech, Palo Alto, CA, USA)"  
5527..9320  
/note="Retroviraler Vektor p5NM"  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.76e-113 Length: 9320  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conserved: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
Gaps: 0  
DB: 6  
US-10-006-922a-12 (1-225) x AK663075 (1-9320)  
Qy 1 MetArgSerSerLySAsnValIleLySglnPheMetArgPheLySValArgMetGluGly 20  
Db 4041 GTGGCTCTCTCCAAAGACGTATCAAGAGATTCAATGCGCTTCAAGTGGCATGAGAGGAC 4100  
Qy 21 ThrValaenGlyHisGlnPheGluIleGluGlyGluGlyAyrProTyrGluGly 40  
Db 4101 ACCGTGAACGGCCACAGATTCAGATTCAGAGCGCGGCGAGGCGCCCTTACAGAGGC 4160  
Qy 41 HisAsnThrValLySLeuLySValThrLySgLyGlyProLeuProPheAlaTrpAspIle 60  
Db 4161 CACCAACCGGTGAACCTGAAGGTGACCAAGGCGGCGCCCTTCCCTTCCGCTGGAGATC 4220  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
Db 4221 CTGTCCCCCACTTCAAGTACGAGCTTCAAGGTGATGATGAGACACCCGCGACATCCCC 4280  
Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpgLuarValMetAsnPheGlu 100  
Db 4281 GACTACAAAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGGCGTATGAACTTCGAG 4340  
Qy 101 AspGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPhelIleTyr 120  
Db 4341 GACGGCGGCGTGTACCGTGAACCAAGATCTCTCTGACAGAGCGCTTCACTAC 4400  
Qy 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140  
Db 4401 AAGGTGAAGTTCATCGGCGTGAATCTTCCCTCGAGCGGCCCGTATGCAAGAAAGACC 4460  
Qy 141 MetGlyTTrpGlnAlaSerThrGlnArgLeuTyrProArgAspGlyValLeuLySgLyGln 160  
Db 4461 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAG 4520  
Qy 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGlnPheLySserIle 180  
Db 4521 ATCCACAAAGCCCTGAAGCTGAAGAGCGGCGCCTACTCTGTGGAGTTCAATGTCATC 4580  
Qy 181 TyrMetAlaIyLySProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAap 200  
Db 4581 TACATGGCCAAAGAGCCCGTGAAGCTGCCGGCTACTACTAGTGAACCTCAAGCTGGAC 4640  
Qy 201 IleThrSerHisaenGluAapTyrThrIleValGluGlnTyrGluAyrGThglGluIyAyr 220  
Db 4641 ATCACTCCCAACAGAGACTACACATCTGGAGCAGTACGAGGACCGAGGGCGCG 4700  
Qy 221 HisHisLeuPheLeu 225  
Db 4701 CACCACTGTTCTTG 4715  
RESULT 20  
AY569780/c  
LOCUS AY569780 12404 bp DNA circular SYN 05-APR-2004

DEFINITION Cloning vector pTCCR-Auto, complete sequence.  
ACCESSION AY569780  
VERSION AY569780.1 GI:45861238  
KEYWORDS  
SOURCE Cloning vector pTCCR-Auto  
ORGNANISM Cloning vector pTCCR-Auto  
other sequences; artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 12404)  
TITLE Wu, G., Wang, Y. and Zhu, Z.  
AUTHORS Direct Submission  
JOURNAL Submitted (09-MAR-2004) Fish Genetics, Institute of Hydrobiology,  
Chinese Academy of Science, 7# South Road of East Lake, Wuhan,  
Hubei 430072, People's Republic of China  
location/Qualifiers  
FEATURES  
source 1..12404  
/organism="Cloning vector pTCCR-Auto"  
/mol\_type="other DNA"  
/db\_xref="taxon:268734"  
419..452  
misc\_feature  
/note="loxP Rg"  
complement(505..555)  
/note="SV40 polyA"  
complement(707..1387)  
/codon\_start=1  
/transl\_table=11  
/product="RFP"  
/protein\_id="AAS78500.1"  
/db\_xref="GI:45861239"  
/translation="MVRSSKNVLEKPMKFKVMEGTVNGHEFEIEEGEGGRPYEGHNT  
VKLKVTKGGLPFPAMDILSPQYGSKYVGHGPAIDPYKLSPEEGKMERVNFED  
GGVAVTVQDSSLQDGCPIYKQFIGVNPSPDGPMQKTMGMEASTERYLDVYLGK  
EIHNAKLKQDGHVLEFKSIYMAKKPQLPGYVYVDSKLDITSHNEDYTIIVEQERT  
EGRHHLPL"  
promoter complement(1447..2036)  
/note="CMV promoter"  
complement(2387..2761)  
/codon\_start=1  
/transl\_table=11  
/product="zeocin"  
/protein\_id="AAS78501.1"  
/db\_xref="GI:45861240"  
/translation="MAKLTSAPVLTARDAVAGAVEFWTDRLGFSRDPVEDDPAGVVRD  
DVTLFISAVQOVVDNTLAWVWVGDELTAWESEVSTFRDASGPAIMEISQPM  
GREFLRDPAGNCVHFAVEBD"  
CDS complement(2860..2990)  
/note="T7 terminator"  
complement(3056..4097)  
/codon\_start=1  
/transl\_table=11  
/product="Cre"  
/protein\_id="AAS78502.1"  
/db\_xref="GI:45861241"  
/translation="MANLLTVHQNLIPALPVDATSDVERKNLMDMFRDROAFSEHTWKV  
LLSVCRSWAACKLNNRKMFPAPBEDVDLLYLDARGLAVKTIQOHLGQNLNHRRS  
GLPRPSDNASLVWRIRIKENVDAGERAKALAFERDFQVSLMENSRCDDIRN  
LAPLGIAVNTLLRIAEIARIKDISRTDGRMLIHGRTKTSLSTAGEKALSLGDT  
KLVRWISVSGVADDPNNYLFCVRKXGVNAPASATSOIETPALGIPATTRLIYGAK  
DDSGCRILASGHSRVCANAPDMARAGVSTEIFIMQAGMTVNVITVNTIRLDBSETGA  
WVRLIEDGD"  
misc\_feature complement(4109..4683)  
/note="internal ribosome entry sequence from  
encephalomyocarditis virus"  
complement(4745..4762)  
/note="T7 promoter"  
complement(4837..4967)  
/note="T7 terminator"  
complement(5009..7660)  
/codon\_start=1  
/transl\_table=11  
/product="T7 RNA polymerase"  
/protein\_id="AAS78503.1"  
/db\_xref="GI:45861242"  
/translation="MNTINIAKNDPSDIELAIPFNTLADHYGRLAREQLALHESY

EMGEARPRKMEROLKAGEVADNNAKPLITLLPKMIARINDMFEVYKAKRGKRP  
FOPLQETKPEAVANITTKTTLACTSLADNTTVOVASAIGATIDBARFSGIRDL  
HFKKVEBOLANKRAGVYTKKAFQVYVADMLSKLLGSEANSHKHSIHVRCITC  
MLIESTWMSLHRQNAVVGQDSEITLAPFPAALITRGALGIGSWFQCVAPPK  
PMVTGITGGGWANGRRLPALVRTSKALMEVEVYKALINIAQNTMKINKV  
LAVANVITKMGHCEVEDIPALIERBELPMKEDIMNBEALPAMRAAAYVRKQAK  
SRISIEGFMLEQANKFANKAIWPFYNNMGRGYAVSMFNPQNDMTKGLTLAKG  
PIGKEGYVWLKIHGANCAGVDKVPFPRIRKIEENHENIMCAKSPLENTWADQSP  
FCPLARCFEYAGVQHGLSYNCSPLAFDSCSGCIQHSAMLDEVGRAVNLDSBT  
VQDTIGIVAKVNSILQADAINGTDNVAVTIDNTSISBKVKLGTRALLAGOMLAVG  
VTRSVTRSVWTLVAVGSKFGRQVLEDITQPAIDSGKGLMTQPNQPAVMAKLW  
ESVSTVVAVEANMMLKSAKLLAASVKOKTGBIILKRCVAMVYDPGPPVQVEYK  
KPIQTRNLMEFLGQFRLQPTINTKDSIEDAHKQESGIAPYFVHSDGSHLRTVVA  
HEKYGISPLILHDSFGTIPADANILKRAVRETVWDVYESCDVLADFYDQADQLHES  
QLDKMPALPAKGNLNLDIIESDPAPA"  
misc\_signal complement(7670..7690)  
/note="SV40 T-antigen NLS"  
complement(7702..8276)  
misc\_feature  
/note="internal ribosome entry sequence from  
encephalomyocarditis virus"  
complement(8338..8355)  
promoter  
/note="T7 promoter"  
complement(8744..9826)  
/codon\_start=1  
/transl\_table=11  
/product="Lact"  
/protein\_id="AAS78504.1"  
/db\_xref="GI:45861243"  
/translation="MKRVTLLDAEYVAGSVQTVSRVYNQASHVSAKREKEVAMAE  
LNTYIPNVVAQOLAKQSLILGVATSSIALHAPSOIVAIKSRADQLASVAVSWERS  
GVEACKSAVNLNLAQRVSGILINPLDDOPIAIEAACTVNPALFLVSDTPYNSII  
FSHEDGRLGVENHVALGHQOIALAPLSVSRLAGNKHKLTLRNOIQPIABREG  
DMSMSGFOOTMOMLNEGIVPTMVLVNDOMLAGMBAITRESGRVADISVGVYDT  
EDSCYIPEPSTTIKQDFRLIGQTSVVDLLQLSQQAVKGNLLPVSILVKRTTLA  
PNTQTASPRALADSLQOLARQVSRLESQ"  
misc\_feature 10113..10146  
/note="loxP LE"  
10585  
/note="Col E1"  
complement(11344..12204)  
/codon\_start=1  
/transl\_table=11  
/product="Amp"  
/protein\_id="AAS78505.1"  
/db\_xref="GI:45861244"  
/translation="MSIQHFRVALIPFPAACLPVPAHPETVLYKXADQLGARVY  
IBLDNSGKILSEFRPERFPWSTFVLLCGAVLSRIDAQEBQGRIRIHSQNDLVE  
YSPTERKHLVDGMTVRELCSAAITMSDNTANLLLTIGGKELTARFLHNGDHTVRL  
DRMEPELNEAIIPNDERDTTPVAMATYLRKLLTGEILLTASRQQLIDMEADKVA  
GLPLRSALPAGWFLADSGAGBERSGRIIALGPDGKPSRIIVLYTTGSAQTYMERRQIA  
EIGASLIKHM"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,24e-112 Length: 12404  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922A-12 (1-225) x AY569780 (1-12404)  
Qy 1 MectArgSerIybaAnvalIlelyGluPhemeTarPhelyValArgMetGluY 20  
Db 1384 GTGGCTCTCCCAAGACGTCATCAAGAGTTCATYGGCGTTCAGAGCGCATGAGGCGC 1325  
Qy 21 ThrValaenglyHsgluPhegluIleGluYgluYgluYgluYArgProTyGluY 40  
Db 1324 ACCGTGAACGCCACGAGTTCGAGATCGAGGCGCAGGCGCGGCGCCCTTACGAGGCGC 1265  
Qy 41 HisaenThrVallyleuLyValIThrlyGluYgluYProleuProPhelaITrpAspIle 60



Score: 1196.00 Matches: 221  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.52% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX348043 (1-666)

QY 5 LysAsnValIleIleValGluPheMetArgPheIysValArgMetGluGlyThrValAsnGly 24  
DB 1 AAGAAATGTTATCAAGAGATTCATGAGGTTTAAAGTTCCGATGGAAGAAACGGTCAATGG 60  
QY 25 H1sgLupheGluIleGluGlyGluGlyArgProTyrGluGlyYhiAsnThrVal 44  
DB 61 CACGAGTTTGAATAGAAAGCGAAGAGAGAGGAGCCATACGAAGCCACATACCGTA 120  
QY 45 LysLeuIysValThrIleGlyGlyProIeuProPheAlaTTPAspIleLeuSerProGln 64  
DB 121 AAGCTTAAGGTAAACCAAGGGGGGACCTTGGCCATTGGCTGGATATTTTGCACACAA 180  
QY 65 PheGlnTyrGlySerIysValTyrValIlyshIleProAlaIleProAspTyrIlyIys 84  
DB 181 TTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCCGACATACCAAGACTATATAAAG 240  
QY 85 LeuSerPheProGluGlyPheIleTrrpGluArgValMetAsnPheGluAspGlyGlyVal 104  
DB 241 CTGTCAATTTCCGAAAGATTTTAAATGGGAAAGGTCATGAACCTTTGAAAGCGGTGGCTC 300  
QY 105 ValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyrIlyValIlyAsp 124  
DB 301 GTTACTGTAAACCAAGATTCACAGTTTGCAAGATGCTGTTCATCTACAGGTCAAGCTTC 360  
QY 125 IleGlyValAsnPheProSerAspGlyProValMetGlnIlyIysThrMetGlyTrrpGlu 144  
DB 361 ATTGGCTGGAACCTTCTCCGATGACCTGTTATGCAAAAGAAACAATGGCTGGGAA 420  
QY 145 AlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGluIleIshIlyshAla 164  
DB 421 GCCAGCACTGAGCGTTTGTATCTCGTATGCGGTGAAAGAGAAATTCATTAAGGCT 480  
QY 165 LeuIlyLeuIysAspGlyGlyYhiIshTyrLeuValGluPheIysSerIleTyrMetAlaIys 184  
DB 481 CTGAAGCTGAAAGACGGTGTCTATCTAGTGAATTCAAAGATTTTAACTAGCCAAAG 540  
QY 185 LysProValGlnLeuProGlyTyrTyrTyrValIleAspSerIysLeuAspIleThrSerHis 204  
DB 541 AAGCTGTGACGCTACAGGAGTACTAGTGTGACTCCAAACTGGATATACAGCCAC 600  
QY 205 AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgYhiIshIleuPhe 224  
DB 601 AACGAAAGCTATACATCGTTGAGAGCATATGAAAGAACCGAGGAGCGCCACATCTGTT 660  
QY 225 Leu 225  
DB 661 CTT 663

RESULT 23  
AX353910 LOCUS 666 bp DNA linear PAT 06-FEB-2002  
DEFINITION Sequence 22 from Patent WO0204509.  
ACCESSION AX353910  
VERSION AX353910.1 GI:1861888  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE  
1 Odenthal, M. and Jung, D.  
Gene expression, genome alteration and reporter expression in  
myofibroblasts and myofibroblast-like cells  
Patent: WO 0204509-A 22 17-JAN-2002;  
Odenhal, Margarete (DB)  
location/Qualifiers

## SOURCE

1. .666  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Red Fluorescent Protein"

## ORIGIN

## Alignment Scores:

Pred. No.: 9.44e-113 Length: 666  
Score: 1196.00 Matches: 221  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.52% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX353910 (1-666)

QY 5 LysAsnValIleIleValGluPheMetArgPheIysValArgMetGluGlyThrValAsnGly 24  
DB 1 AAGAAATGTTATCAAGAGATTCATGAGGTTTAAAGTTCCGATGGAAGAAACGGTCAATGG 60  
QY 25 H1sgLupheGluIleGluGlyGluGlyArgProTyrGluGlyYhiAsnThrVal 44  
DB 61 CACGAGTTTGAATAGAAAGCGAAGAGAGAGGAGGCCATACGAAGCCACAAATACCGTA 120  
QY 45 LysLeuIysValThrIleGlyGlyProIeuProPheAlaTTPAspIleLeuSerProGln 64  
DB 121 AAGCTTAAGGTAAACCAAGGGGGGACCTTGGCCATTGGCTGGATATTTTGCACACAA 180  
QY 65 PheGlnTyrGlySerIysValTyrValIlyshIleProAlaIleProAspTyrIlyIys 84  
DB 181 TTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCCGACATACCAAGACTATATAAAG 240  
QY 85 LeuSerPheProGluGlyPheIleTrrpGluArgValMetAsnPheGluAspGlyGlyVal 104  
DB 241 CTGTCAATTTCCGAAAGATTTTAAATGGGAAAGGTCATGAACCTTTGAAAGCGGTGGCTC 300  
QY 105 ValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyrIlyValIlyAsp 124  
DB 301 GTTACTGTAAACCAAGATTCACAGTTTGCAAGATGCTGTTCATCTACAGGTCAAGCTTC 360  
QY 125 IleGlyValAsnPheProSerAspGlyProValMetGlnIlyIysThrMetGlyTrrpGlu 144  
DB 361 ATTGGCTGGAACCTTCTCCGATGACCTGTTATGCAAAAGAAACAATGGCTGGGAA 420  
QY 145 AlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGluIleIshIlyshAla 164  
DB 421 GCCAGCACTGAGCGTTTGTATCTCGTATGCGGTGAAAGAGAAATTCATTAAGGCT 480  
QY 165 LeuIysLeuIysAspGlyGlyYhiIshTyrLeuValGluPheIysSerIleTyrMetAlaIys 184  
DB 481 CTGAAGCTGAAAGACGGTGTCTATCTAGTGAATTCAAAGATTTTAACTAGCCAAAG 540  
QY 185 LysProValGlnLeuProGlyTyrTyrTyrValIleAspSerIysLeuAspIleThrSerHis 204  
DB 541 AAGCTGTGACGCTACAGGAGTACTAGTGTGACTCCAAACTGGATATACAGCCAC 600  
QY 205 AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgYhiIshIleuPhe 224  
DB 601 AACGAAAGCTATACATCGTTGAGAGCATATGAAAGAACCGAGGAGCGCCACATCTGTT 660  
QY 225 Leu 225  
DB 661 CTT 663

RESULT 24  
AY679107 LOCUS 711 bp mRNA linear INV 25-AUG-2004  
DEFINITION Discoma sp. RC-2004 enhanced red fluorescent protein R+ mRNA,  
complete cds.  
ACCESSION AY679107  
VERSION AY679107.1 GI:51472046  
KEYWORDS

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
SOURCE  
CDS

Discosoma sp. RC-2004  
Discosoma sp. RC-2004  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
1 (bases 1 to 711)  
Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.  
Cloning of Anthozoan Fluorescent Protein Genes  
Unpublished  
2 (bases 1 to 711)  
Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.  
Direct Submission  
Submitted (06-JUL-2004) MBF, University of Miami, 4600 Rickenbacker  
Cswy, Miami, FL 33129, USA  
Location/Qualifiers  
1..711  
/organism="Discosoma sp. RC-2004"  
/mol\_type="mRNA"  
/db\_xref="taxon:289055"  
1..711  
/note="GFP-like; DsRed+"  
/codon\_start=1  
/product="enhanced red fluorescent protein R+"  
/protein\_id="AAU04444.1"  
/db\_xref="GI:51472047"  
/translation="MSCSKNYLKEFMRKRVNMGTVNGHEBEIEGEGRPYGHNTV  
KLKVTGSPPLPAMDILSPQYSGSKYVVKPADIPDYKULSPFGFKMERVMNEDG  
GVVTVTQDSIQDGGCFITKVKFIGNVPSDGFVWQKTMGWEASTERYLPDGVLKGE  
IHKLKLDQGHYIVFRTIYVAKKRPVLPGYVYDSKLDITSHNKDYTIYEQYRTE  
GRHHLFLKAEIGSNVGR"

ORIGIN

Alignment Scores:  
Pred. No.: 1,02e-112 Length: 711  
Score: 1196.00 Matches: 221  
Percent Similarity: 99.11% Conservative: 2  
Best Local Similarity: 98.22% Mismatches: 2  
Query Match: 98.52% Indels: 0  
DB: 2 Gaps: 0

US-10-006-922A-12 (1-225) x AY679107 (1-711)

1 MetArgSerSerLyAsnValIleLySGIuPhMeTarGPhelyValArgMetGluGly 20  
1 ATGAGTTGTTCCAAGATGTTATCAAGAGTTCATAGGTTTAAAGTTGTAAGAGCA 60  
21 ThrValAsnGlyVHIGluPhGluIleGluGluGluGluGluGluGluGluGluGlu 40  
61 ACGGTCAATGGGCAACAGTTTGAAATAGAAAGCGAAAGGAGGAGGACATACGAAGGC 120  
41 HisAsnThrValIleLeuLeuValThrLySGIyGlyProLeuProPheAlaTrpAspIle 60  
121 CACATACCGCTAAAGCTTAAGTACCAAGGGGAGCCTTTCGCACTTCTGGGATATT 180  
61 LeuSerProGlnPheGlnIleGlySerLySerLyValTyrValIleHisProAlaAspIlePro 80  
181 TTGTCAACCAATTTCAATATGGAACCAAGTATATGCAAGCATCTCCGACATACCA 240  
81 AspTyrLyLeuLeuSerPheProGluGlyPheLySerLyValIleValMetAsnPhGlu 100  
241 GACTATAAAGAGCTGTCATTTCTGAAGAGATTAAATGGAAGAGGTCATGAACTTTGAA 300  
101 AAGGlyGlyValValThrValThrGlnAAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
301 GACGGTGGGCTCTTACTGTAACCAAGATTCAGTTTGAGAGATGGCTGTTTCATCTAC 360  
121 LyValAllyPheIleGlyValAlaAsnPhProSerAspGlyProValMetGlnLySerThr 140  
361 AAGGTCAAGTTCAATGGCGTGAACCTTCTCTGATGAGACCTGTATGCAAAAGAAAGACA 420  
141 MetGlyTyrGlnAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGIyGlu 160  
421 ATGGGCTGGGAACCGACACTGAGCGTTGTATCTCTGATGCGGCTGTTGAAAGAGAG 480

QY 161 IleHisValAlaLeuLyLeuLyAspGlyGlyHisTyrLeuValGluPhelySerIle 180  
DB 481 ATTATTAAGGCTGGAAGTTGAAGACGGTGGTCACTTACTAGTGAATTCAAAACTATT 540  
QY 181 TyrMetAlaLyLyProValGlnLeuProGlyTyrTyrTyrValAlaAspSerLyLeuAsp 200  
DB 541 TACATGGCAAGAAAGCTGTGCACCTACCGAGGAGTACTATGTTGACTCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnIleTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAACAAGACTATACATCGTTGAGCAGTATGAAGAACCGAGGAGAGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCTT 675

RESULT 25  
AF545828  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
SOURCE  
CDS

AF545828 678 bp mRNA linear INV 04-JUN-2004  
Discosoma sp. JW-2002 orange fluorescent protein FP586 mRNA,  
complete cds.  
AF545828  
AF545828.1 GI:3333763  
Discosoma sp. JW-2002  
Discosoma sp. JW-2002  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
1 (bases 1 to 678)  
Wiedemann, J. and Glrod, A.  
Direct Submission  
Submitted (12-SEP-2002) Abteilung Allgemeine Zoologie und  
Endokrinologie, Universitaet Ulm, Albert Einstein Allee 11, Ulm  
89069, Germany  
Location/Qualifiers  
1..678  
/organism="Discosoma sp. JW-2002"  
/mol\_type="mRNA"  
/db\_xref="taxon:208461"  
/note="Isolated from specimen with smooth disc"  
1..678  
/note="GFP-like protein, orange-red; dsFP586; similar to  
fluorescent protein drFP583 (dsRed)"  
/codon\_start=1  
/product="orange fluorescent protein FP586"  
/protein\_id="AAQ11987.1"  
/db\_xref="GI:3333764"  
/translation="MSCSKNYLKEFMRKRVNMGTVNGHEBEIEGEGRPYGHNTV  
KLKVTGSPPLPAMDILSPQYSGSKYVVKPADIPDYKULSPFGFKMERVMNEDG  
GVVTVTQDSIQDGGCFITKVKFIGNVPSDGFVWQKTMGWEASTERYLPDGVLKGE  
IHKLKLDQGHYIVFRTIYVAKKRPVLPGYVYDSKLDITSHNKDYTIYEQYRTE  
GRHHLFL"

ORIGIN

Alignment Scores:  
Pred. No.: 1,95e-112 Length: 678  
Score: 1193.00 Matches: 221  
Percent Similarity: 98.67% Conservative: 1  
Best Local Similarity: 98.22% Mismatches: 3  
Query Match: 98.27% Indels: 0  
DB: 2 Gaps: 0

US-10-006-922A-12 (1-225) x AF545828 (1-678)

1 MetArgSerSerLyAsnValIleLySGIuPhMeTarGPhelyValArgMetGluGly 20  
1 ATGAGTTGTTCCAAGATGTTATCAAGAGTTCATAGGTTTAAAGTTGTAAGAGCA 60  
21 ThrValAsnGlyVHIGluPhGluIleGluGluGluGluGluGluGluGluGluGlu 40  
61 ACGGTCAATGGGCAACAGTTTGAAATAGAAAGCGAAAGGAGGACATACGAAGGC 120  
41 HisAsnThrValIleLeuLeuValThrLySGIyGlyProLeuProPheAlaTrpAspIle 60

```

Db      121 CACAAATACCGTAACGTTAAAGGTAACCAAGGGGGACCTTCCCATTTGGGATATT 180
Qy      61 LeuSerProGlnPheGlnIlyrGlySerIyValIlyrValIyHisProAlaAspIlePro 80
Db      181 TTGTCAACCAACATTTTCAGTATGGAAGCAAGGTATATGTCACACCCCTGCCACATACCA 240
Qy      81 AspTyrIlyValIyLeuSerPheProGluGlyPheIlySTPGLuArgValIyMetAsnPheGlu 100
Db      241 GACTTATTAAGGCTGTCACTTTCTCTTAAGGATTTAAATGGGAAAGGGTCAATGAACCTTTGAA 300
Qy      101 AspGlyGlyValIyValIyThrValIyThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      301 GACGGGTGGGCTGTACTGTAACTGTAACCAAGATTCAGATTGGACAGATGGCTGTTTCACTTAC 360
Qy      121 IyValIyIyPheIleGlyValIyAsnPheProSerAspGlyProValIyMetGlnIyIySerIlyThr 140
Db      361 AAGGTCAAGTTCAATGGGGGTGAACCTTCTCTGTATGACCTGTATGCAAAAGAAAGACA 420
Qy      141 MetGlyTTPGLuAlaSerThrGluArgLeuTyrProArgAspGlyValIleuIyGlyIy 160
Db      421 ATGGCTGGGAGAACCAACGACTGAGCGTTTGTATCTCTGTATGGCGCTGTTGAAAGAGAT 480
Qy      161 IlyHisIyValIyLeuIyLeuIyAspGlyGlyIyHisTyrLeuValIyGluPheIySerIle 180
Db      481 ATTCAATAAGGCTCTGAGAGCTGAAGACGGTGTCTATTAAGTTGAATTCAAAAGTATTT 540
Qy      181 TyrMetAlaIyIyIyPheProValIyGlnIyLeuProGlyTyrTyrTyrValIyAspSerIyLeuAsp 200
Db      541 TACATGGCAAAAGAACCTGTGTGCACTACCAAGGGTCTCTATGTATGACTCCCAACCTGAT 600
Qy      201 IlyThrSerHisAsnGluAspTyrThrIleValIyGluGlnIyThrGluArgThrGluIyArg 220
Db      601 ATTAACAAGCCCAACGAAGACTATACATCGTTGACGATGAAAGAACGAGGGAAGCC 660
Qy      221 HisHisLeuPheLeu 225
Db      661 CACCATCTGTTCTT 675

RESULT 26
LOCUS      AX370408                678 bp    DNA        linear    PAT 16-FEB-2002
DEFINITION Sequence 5 from Patent WO0196373.
ACCESSION  AX370408
VERSION     AX370408.1  GI:18857492
KEYWORDS    .
SOURCE      synthetic construct
            synthetic construct
            other sequences; artificial sequences.
ORGANISM    Pradkov, A. P. and Terekikh, A.
REFERENCE   1
AUTHORS     Fluorescent timer proteins and methods for their use
TITLE       Patent: WO 0196373-A 5 20-DEC-2001;
JOURNAL     Clontech Laboratories Inc. (US)
FEATURES    location/Qualifiers
            1..678
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="variant of sequence from Discosoma sp."

ORIGIN
Alignment Scores:
Pred. No.:      3,13e-112      Length:      678
Score:          1191.00      Matches:      220
Percent Similarity: 98.67%      Conservative: 2
Best Local Similarity: 97.78%      Mismatches: 3
Query Match:    98.11%      Indels:      0
DB:             6      Gaps:      0

US-10-006-922a-12 (1-225) x AX370408 (1-678)
Qy      1 MetArgSerSerIyAsnValIlyIyGluPheMetArgPheIyValIyArgMetGluIy 20

```

```

Db      1 ATGGGCTCTCCGAGAAAGTCATACCGAGTTCATGGCTTCAAGGTGCGCATGAGGGC 60
Qy      21 ThrValAsnGlyHisGluPheGluIlyIyGluIyGluIyArgProTyrGlyIy 40
Db      61 ACCGTGAACCGGCCACGAGTTGATGATGAGGGCCAGGGCCGAGCCGCTACGAGGGC 120
Qy      41 HisAsnThrValIyIyLeuIyValIyThrIyArgIyGlyIyProIyProPheAlaIyTyrAspIle 60
Db      121 CACAAACCGTAAGCTGAAAGGTGACCAAGGGCGGCCCCCTTGCTTGCCTGGGACATC 180
Qy      61 LeuSerProGlnPheGlnIlyrGlySerIyValIyValIyHisProAlaAspIlePro 80
Db      181 CTGTCCGCCCACTTCCAGTACGCTCCCAAGGTGACGAAAGAACCCCGCGCATATCCC 240
Qy      81 AspTyrIlyValIyLeuSerPheProGluGlyPheIlySTPGLuArgValIyMetAsnPheGlu 100
Db      241 GACTAACAAGAACGTGTCTCTCCCGAAGGGCTTCAAGTGGGAGCGCGTGAATGAACCTTCAG 300
Qy      101 AspGlyGlyValIyValIyThrValIyThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      301 GACGGCGGCTGGGACCGTGACCAAGACTCTCTCTGACAGAGCGCTGCTTCACTTAC 360
Qy      121 IyValIyIyPheIleGlyValIyAsnPheProSerAspGlyProValIyMetGlnIyIySerIlyThr 140
Db      361 AAGGTCAAGTTCACTGGGCGTGAATCTTCCCTCCGACGGCCCCCGATGACAAAGAACACC 420
Qy      141 MetGlyTTPGLuAlaSerThrGluArgLeuTyrProArgAspGlyValIleuIyGlyIy 160
Db      421 ATGGCTGGGAGGGCTCTCACCGAGCGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 480
Qy      161 IlyHisIyValIyLeuIyLeuIyAspGlyGlyIyHisTyrLeuValIyGluPheIySerIle 180
Db      481 ATCCAAAGGCCCTGAAGCTGAAAGACGGGCGCCACTACCTGTGTGAGTTCAATCTCATC 540
Qy      181 TyrMetAlaIyIyIyPheProValIyGlnIyLeuProGlyTyrTyrTyrValIyAspSerIyLeuAsp 200
Db      541 TACATGGCAAAAGAACCGCGTGAAGCTGAGGCGCCACTACCTGTGTGAGTTCAATCTCATC 600
Qy      201 IlyThrSerHisAsnGluAspTyrThrIleValIyGluGlnIyThrGluArgThrGluIyArg 220
Db      601 ATCACTCTCCCAACAAGGACTACCATCGTGAAGACGATGACGACCGACCGAGGGCGCG 660
Qy      221 HisHisLeuPheLeu 225
Db      661 CACCATCTGTTCTT 675

RESULT 27
LOCUS      AX824732                678 bp    DNA        linear    PAT 11-DEC-2003
DEFINITION Sequence 14 from Patent WO02068459.
ACCESSION  AX824732
VERSION     AX824732.1  GI:39750595
KEYWORDS    .
SOURCE      synthetic construct
            synthetic construct
            other sequences; artificial sequences.
ORGANISM    Non aggregating fluorescent proteins and methods for using the same
REFERENCE   1
AUTHORS     Patent: WO 02068459-A 14 06-SEP-2002;
TITLE       Clontech Laboratories Inc. (US)
JOURNAL     Clontech Laboratories Inc. (US)
FEATURES    location/Qualifiers
            1..678
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="non-aggregating mutant"

ORIGIN
Alignment Scores:
Pred. No.:      3,13e-112      Length:      678
Score:          1191.00      Matches:      220
Percent Similarity: 98.67%      Conservative: 2

```



Best Local Similarity: 97.78% Mismatches: 3  
 Query Match: 98.11% Indels: 0  
 DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AX824732 (1-678)

```

OY 1 MetArgSerSerLyAsnValIleLygLuPhMetArgPheLyValArgMetGluGly 20
DB 1 ATGGCTCTCCGAGAAAGTTCATCAGGATTCAGCGCTTCAGAGTGCATGAGAGGCG 60
OY 21 ThrValaAngLyHLeuLyPheGluIleGluGlyGluGlyArgProTyArgGluGly 40
DB 61 ACCGGGAACGGCCAGAGTTCAGAGTTCAGAGGCGGCGGCGGCGGCGGCGGCGGCG 120
OY 41 HisAsnThrValLyLeuLyValThryLygLyGlyProLeuProPheAlaTyrAspIle 60
DB 121 CACAAACCGTGAAGTTCAGAGTTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCG 180
OY 61 LeuSerProGlnPheGlnTyArgLySerLyValTyValLyHisProAlaAspIlePro 80
DB 181 CTGTCCCGCCAGTTCAGAGTTCAGAGGCGTTCAGAGGCGGCGGCGGCGGCGGCGGCG 240
OY 81 AspTyArgLyLeuSerPheProGluGlyPheLySTPGLuArgValMetAsnPheGlu 100
DB 241 GACTACAAAGAAAGCTCTCTCCCGAGGCGTTCAGAGTTCAGAGGCGGCGGCGGCGGCG 300
OY 101 ArgGlyLyValValThryValThrgLnAspSerSerLeuGlnAspGlyCyAspHeIleTy 120
DB 301 GACGGCGGGGTCGCAACCTGTACCCAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
OY 121 LySValLyPheIleGlyValaAsnPheProSerAspGlyProValMetGlnLyLyThr 140
DB 361 AAGGGAAGTTCATGGGGGTGAATCTCCCTCCGACGGGCGGCGTGAAGAAAGAAC 420
OY 141 MetGlyTTPGLuAlaSerThrgLnArgLeuTyTProArgAspGlyValaLeuLygLyGlu 160
DB 421 ATGGCTGGGAGGCTCTCCAGCGGCGCTGTACCCCGGAGCGGCGTGAAGAGCGAG 480
OY 161 IleHisLyValaLeuLyLeuLyAspGlyLyHisTyLeuValaGluPheLySerIle 180
DB 481 ATCCCAAGGCGCTGAAGCTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
OY 181 TyrMetAlaLyLyPheProValaGlnLeuProGlyTyTyrTyValaAspSerLyLeuAsp 200
DB 541 TACATGGCCAAAGAGCGGTGAGCTGCCGGCTACTACTAGTGAACCAAGCTGGAGC 600
OY 201 IleThrSerHisAsnGluAspTyThryIleValaGluGlnTyTArgGluArgThrgLnArg 220
DB 601 ATCACCCTCCCAACAGAGACTACCATCTGAGAGCAAGTACGAGGCGACCGAGGGCGC 660
OY 221 HisHisLeuPheLeu 225
DB 661 CACCACTGTCTCTG 675

```

RESULT 28  
 LOCUS AY679106 921 bp mRNA linear INV 25-AUG-2004  
 DEFINITION Discosoma sp. RC-2004 red fluorescent protein R1 mRNA, complete  
 cde.  
 ACCESSION AY679106  
 VERSION AY679106.1 GI:51472044  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Discosoma sp. RC-2004  
 Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
 Corallimorpharia; Discosomatidae; Discosoma.  
 1 (bases 1 to 921)  
 2 (bases 1 to 921)  
 TITLE Cloning of Anthozoan Fluorescent Protein Genes  
 AUTHORS Carter R.W., Gibbs P.D.L. and Schmale M.C.  
 REFERENCE 2 (bases 1 to 921)  
 JOURNAL Unpublished  
 REFERENCES  
 AUTHORS Carter R.W., Gibbs P.D.L. and Schmale M.C.  
 TITLE Direct Submission

JOURNAL Submitted (06-JUL-2004) MBF, University of Miami, 4600 Rickenbacker  
 Cswy, Miami, FL 33129, USA  
 FEATURES  
 source  
 1..921  
 /organism="Discosoma sp. RC-2004"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:289055"

5'UTR  
 CDS  
 1..66  
 67..777  
 /note="GFP-like; DsRed1"  
 /codon\_start=1  
 /product="red fluorescent protein R1"  
 /protein\_id="AAU04443.1"  
 /db\_xref="GI:51472045"

/translation="MSSKKNVYKFPKPKRMESTVNGHREIREGEGRPYEGHNTY  
 KLKVTGGPLPFPAMDILSPQVGSKYVYKHPALIPYKLSFPFGKRWKRVNFPEDG  
 GVVTVDPSLDGCFYKVKFLGVNPSDPSVQKTKMTGMEASTELYPYDGLKGR  
 IHKALKLDGGHYLVEFKTYMAKPKVQLPRYYVDSKLDITSHKDYTYVEYERTE  
 GRHHLFLKELSGSVNGER"  
 778..921  
 882..887

3'UTR  
 polyA\_signal  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4,54e-112 Length: 921  
 Score: 1191.00 Matches: 220  
 Percent Similarity: 98.67% Conservative: 2  
 Best Local Similarity: 97.78% Mismatches: 3  
 Query Match: 98.11% Indels: 0  
 DB: 2 Gaps: 0

US-10-006-922a-12 (1-225) x AY679106 (1-921)

```

OY 1 MetArgSerSerLyAsnValIleLygLuPhMetArgPheLyValArgMetGluGly 20
DB 67 ATGAGTTTCCAAAGATGTTATCAAGAGATTCAAGAGTTCATGAGTTCATGAGAGAG 126
OY 21 ThrValaAngLyHLeuLyPheGluIleGluGlyGluGlyArgProTyArgGluGly 40
DB 127 ACGGTCAATGGGCACGAGATTGAATAGAAGCGGAAGAGAGAGGAGCCATACAGAGC 186
OY 41 HisAsnThrValLyLeuLyValThryLygLyGlyProLeuProPheAlaTyrAspIle 60
DB 187 CACAAATACCGTAACCTTAAGGTAAACAAAGGGGAGACTTGGCCATTGCTTGGGATATT 246
OY 61 LeuSerProGlnPheGlnTyArgLySerLyValTyValLyHisProAlaAspIlePro 80
DB 247 TTGTACACCAATTTCAGTATGAAGCAAGTATGTCAAGCATCTGCCAGCATATCCA 306
OY 81 AspTyArgLyLeuSerPheProGluGlyPheLySTPGLuArgValMetAsnPheGlu 100
DB 307 GACTATTAAGAGCTGTCAATTCCTAGAGATTAAATGGGAAAGGCTCATGAACCTTGAA 366
OY 101 ArgGlyLyValValThryValThrgLnAspSerSerLeuGlnAspGlyCyAspHeIleTy 120
DB 367 GACGGTGGGCTGTACTGTAAACCAAGATCCCAAGTTTGCAGAGATGTTGTTTCAATTAC 426
OY 121 LySValLyPheIleGlyValaAsnPheProSerAspGlyProValMetGlnLyLyThr 140
DB 427 AAGGTCAAGTTCATTTGGCGTGAATCTTCTCTGTATGAGACTGTATGCAAAAGAGACA 486
OY 141 MetGlyTTPGLuAlaSerThrgLnArgLeuTyTProArgAspGlyValaLeuLygLyGlu 160
DB 487 ATGGCTGGGAGCAAGCACTGAGCGTGTGTATCTCGTAGAGCGGTGTTGAAGAGAGG 546
OY 161 IleHisLyValaLeuLyLeuLyAspGlyLyHisTyLeuValaGluPheLySerIle 180
DB 547 ATTCATAAGGCTCTGAAGTTCAGAGAGCGGTGCTACTTACTTGAATTCATAACTATT 606
OY 181 TyrMetAlaLyLyPheProValaGlnLeuProGlyTyTyrTyValaAspSerLyLeuAsp 200
DB 607 TACATGGCAAAAGAGCTGTGACGAGTACAGGAGTACTACTATGTGACTCCAAAGTGAAT 666

```

QY 201 ILeThSerHisAenGluAapPtyrThrIleValGluGlnIlyrGluArgThrgluGlyArg 220  
DB 667 ATAAACACCCACCAACAAAGACTATATCAATCGTTGACGATATGAAGAAGAGGAGGACGC 726  
QY 221 H1SH1SLeuPheLeu 225  
DB 727 CACCATCTGTCCTT 741  
RESULT 29  
AX824731 675 bp DNA linear PAT 11-DEC-2003  
LOCUS AX824731  
DEFINITION Sequence 13 from Patent WO02068459.  
ACCESSION AX824731  
VERSION AX824731.1 GI:39750594  
KEYWORDS  
ORGANISM synthetic construct  
SOURCE synthetic construct  
REFERENCE 1 other sequences; artificial sequences.  
AUTHORS  
TITLE Non aggregating fluorescent proteins and methods for using the same  
JOURNAL Patent: WO 02068459-A 13 06-Sep-2002;  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..675  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="non-aggregating mutant"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,01e-111 Length: 675  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
Gaps: 0  
DB: 6  
US-10-006-922a-12 (1-225) x AX824731 (1-675)  
QY 1 MetArgSerSerLySAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20  
DB 1 ATGGCTCTCCCGAAGACGTATCAACGAGTTCATGCGCTTCAAGTGCAGATGAGGGC 60  
QY 21 ThrValaenGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40  
DB 61 ACCGTGAACGGCCAGATTCAGATTCAGAGGCGAGGCGCGGCCCTTACAGAGGC 120  
QY 41 HisAenThrValLyLeuLyValThrLyGlyGlyProLeuProPheAlaTrrAspIle 60  
DB 121 CACAAACCGTAAGCTGAAGTGAACAAAGGCGGCGGCCCTTGCCTTGGGACATC 180  
QY 61 LeuSerProGlnPheGlnIlyrGlySerLySValIlyrValLyHisProAlaAspIlePro 80  
DB 181 CTGTCCCCCGATTCCAGTACGCGTCCAAAGGTGTCGTGAAGACCCCGCGACATCCCC 240  
QY 81 AspTyrlLySlyLeuSerPheProGluGlyPheLySTrPGluArgValMetAsnPheGlu 100  
DB 241 GACTACAAAGAGCTGTCTTCCCGAAGGCTTCAAGTGGAGCGGTGATGAACTTCAG 300  
QY 101 AspGlyGlyValIValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTy 120  
DB 301 GACGGCGGCGTGGCAACGTGACCAAGATCTCTCCGTGCAAGACGGCTGCTTCACTAC 360  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140  
DB 361 AAGGTGAAGTTCATGCGGTGAATCTTCCCTCCAGCGGCGCGGTGATGCAAGAAAGACC 420  
QY 141 MetGlyTrPGluAlaSerThrGluArgLeuTyTrProArgAspGlyValLeuLySlyGlu 160  
DB 421 ATGGCTGGAGGCGCTTCAACGAGGCGCTGTACCCCGAGCGGCTGTCTGAAGGGCAG 480  
QY 161 IleHisLySAlaLeuLyLeuLyAspGlyGlyHisTyrlLeuValGluPheLySerIle 180

DB 481 ACCCAAGGCGCTGAAGCTGAAGAGCGGCGCACACTACCTGGTGAAGTCAAGTCCATC 540  
QY 181 TyrMetAlaLySlyProValGlnLeuProGlyTyTrTyTrValAspSerLySLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCGTCGACCTGCCGCTTACTTACGTGAGCGCAAGCTGAC 600  
QY 201 ILeThSerHisAenGluAapPtyrThrIleValGluGlnIlyrGluArgThrgluGlyArg 220  
DB 601 ATCACTTCCCAACAGAGACTACACCATGTGAGAGTGAAGCGCACCGAGGCGCGC 660  
QY 221 H1SH1SLeuPheLeu 225  
DB 661 CACCACTGTTCCTG 675  
RESULT 30  
AX666133 1050 bp DNA linear PAT 26-MAR-2003  
LOCUS AX666133  
DEFINITION Sequence 7 from Patent WO02060941.  
ACCESSION AX666133  
VERSION AX666133.1 GI:29290961  
KEYWORDS  
ORGANISM unidentified  
SOURCE unidentified  
REFERENCE 1 unclassified sequences.  
AUTHORS Zhao, M., Xu, M., Jiang, P. and Yang, M.  
TITLE Fluorescent proteins  
JOURNAL Patent: WO 02060941-A 7 08-AUG-2002;  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..1050  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="Coral"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.73e-111 Length: 1050  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
Gaps: 0  
DB: 6  
US-10-006-922a-12 (1-225) x AX666133 (1-1050)  
QY 1 MetArgSerSerLySAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20  
DB 289 ATGGCTCTCCCGAAGACGTATCAACGAGTTCATGCGCTTCAAGTGCAGATGAGGGC 348  
QY 21 ThrValaenGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40  
DB 349 ACCGTGAACGGCCAGATTCAGATTCAGAGGCGAGGCGCGGCCCTTACAGAGGC 408  
QY 41 HisAenThrValLySlyLeuLyValThrLyGlyGlyProLeuProPheAlaTrrAspIle 60  
DB 409 CACAAACCGTAAGCTGAAGTGAACAAAGGCGGCGGCCCTTGCCTTGGGACATC 468  
QY 41 AspTyrlLySlyLeuSerPheProGluGlyPheLySTrPGluArgValMetAsnPheGlu 80  
DB 469 CTGTCCCCCGATTCCAGTACGCGTCCAAAGTGAAGTGAAGACCCCGCGACATCCCC 528  
QY 61 LeuSerProGlnPheGlnIlyrGlySerLySValIlyrValLyHisProAlaAspIlePro 100  
DB 529 GACTACAAAGAGCTGTCTTCCCGAAGGCTTCAAGTGGAGCGCGGTGATGAACTTCAG 588  
QY 101 AspGlyGlyValIValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTy 120  
DB 589 GACGGCGGCGTGGCAACGTGACCAAGATCTCTCCGTGCAAGACGGCTGCTTCACTAC 648  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140

```

Db      649 AAGGGAAGTTCATCGCGGTGAATCTCCCTCCGACGGCCCCCGTATCAGAAAGAACCC
Qy      141 MetGlyTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuysGly 160
Db      709 ATGGCTGGAGGCTTCACCGAGCGCTGTACCCCGGACGGCGTGTGAAGGCGAG
Qy      161 IleHisValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
Db      769 ACCCAAGAGGCTTAAAGCTGAAGAGACGGCGGCACTACCTGTGTGAAGTTCATC
Qy      181 TyrMetAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 200
Db      829 TACATGGCCAAAGAACCCGTGACAGCTGCCGGCTACTACTACGTGACCGCAAGCTGAC
Qy      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluArg 220
Db      889 ATCACCCTCCCAAGAGAGACTACACCATGTGAGAGCACTAGAGGACCGAGGAGCCG
Qy      221 HisHisLeuPheLeu 225
Db      949 CACCACTGTTCTCTG 963

RESULT 31
CO849509      4488 bp      DNA      linear      PAT 23-AUG-2004
DEFINITION    Sequence 6 from Patent WO2004067751.
ACCESSION     CO849509
VERSION       CO849509.1 GI:51507513
KEYWORDS
SOURCE        synthetic construct
ORGANISM      other sequences; artificial sequences.
REFERENCE     1 Couillard-Despres, S., Karl, C., Kuhn, H. G. and Aigner, L.
AUTHORS       Use of regulatory sequences for specific, transient expression
TITLE          in neuronal determined cells
JOURNAL        Patent: WO 2004067751-A 6 12-AUG-2004;
               Klinikum der Universitaet Regensburg (DB)
FEATURES
source        1..4488
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="fusion construct homo sapiens and DsRed2"
ORIGIN
Alignment Scores:
Pred. No.:      1,01e-110      Length:      4488
Score:          1186.00      Matches:      219
Percent Similarity: 98.22%      Conservative: 2
Best Local Similarity: 97.33%      Mismatches: 4
Query Match:    97.69%      Indels:      0
DB:             Gaps:      0
US-10-006-922a-12 (1-225) x CO849509 (1-4488)
Qy      1 MetArgSerSerIysAsnValIleLeuGluPheMetArgPheLeuValArgMetGlu 20
Db      3572 ATGGCCCTCCCTCCAGAACGTATCATCCGAGTTCATGCGCTTCAAGGGCGCATGAGGCGC 3631
Qy      21 ThrValAsnGlyHisGluPheGluIleGluGlyValGluGlyArgProTyrGluGly 40
Db      3632 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCGCCCTCAAGAGGCGC 3691
Qy      41 HisAsnThrValIysLeuLeuValThrLysGlyValProLeuProPheAlaTPAPhile 60
Db      3692 CACAAACACCGTAAGTGAAGTGAACCAAGGGCGGCGCCCTGCGCTGCGGACATC 3751
Qy      61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80
Db      3752 CTGTCCCGCCAGTTCAGTACGGCTCCAAAGGTGATCGTAAGCAACCCCGCGACATCCCC 3811

```

```

Qy      81 AspTyrIysLeuSerPheProGluGlyPheYstrPGluArgValMetAsnPheGlu 100
Db      3812 GACTTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCCCGTATGACTTGAG 3871
Qy      101 AspGlyGlyValIleThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      3872 GACGGCGCGTGGGAGCGGTGACCCAGACTCTCCCTCGACGAGCGGCTGCTTCACTAC 3931
Qy      121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIleuysThr 140
Db      3932 AAGGTGAAGTTCATCGCGGTGAATCTCCCTCCGACGGCCCCGTGATGACAAAGAACCC 3991
Qy      141 MetGlyTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuysGly 160
Db      3992 ATGGCTGGAGGCTTCACCGAGCGCTGTACCCCGGACGGCGTGTGAAGGCGAG 4051
Qy      161 IleHisValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
Db      4052 ACCCAAGAGGCTTAAAGCTGAAGAGACGGCGGCACTACCTGTGTGAAGTTCATC 4111
Qy      181 TyrMetAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 200
Db      4112 TACATGGCCAAAGAACCCGTGACAGCTGCCGGCTACTACTACGTGACCGCAAGCTGAC 4171
Qy      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluArg 220
Db      4172 ATCACCCTCCCAAGAGAGACTACACCATGTGAGAGCACTAGAGGACCGACCGAGGCGCC 4231
Qy      221 HisHisLeuPheLeu 225
Db      4232 CACCACTGTTCTCTG 4246

RESULT 32
AY818375      4546 bp      DNA      circular SYN 12-APR-2005
LOCUS         AY818375
DEFINITION    Cloning vector pSAT6-DsRed2-C1, complete sequence.
ACCESSION     AY818375
VERSION       AY818375.1 GI:56553574
KEYWORDS
SOURCE        Cloning vector pSAT6-DsRed2-C1
ORGANISM      Cloning vector pSAT6-DsRed2-C1
               other sequences; artificial sequences; vectors.
REFERENCE     1 (bases 1 to 4546)
AUTHORS       Tzfira, T., Tian, G. W., Lacroix, B., Vyas, S., Li, J., Leitner-Dagan, Y.,
               Krichavsky, A., Taylor, T., Vainstein, A. and Citovsky, V.
TITLE          pSAT vectors: a modular series of plasmids for autofluorescent
               protein tagging and expression of multiple genes in plants
JOURNAL        Plant Mol. Biol. 57 (4), 503-516 (2005)
PUBMED        15821977
REFERENCE     2 (bases 1 to 4546)
AUTHORS       Tzfira, T., Lacroix, B. and Citovsky, V.
TITLE          Direct Substitution
JOURNAL        Submitted (05-NOV-2004) Department of Biochemistry and Cell
               Biology, State University of New York Stony Brook, 100 Nicolls Rd,
               Stony Brook, NY 11794, USA
FEATURES
source        1..4546
               /organism="Cloning vector pSAT6-DsRed2-C1"
               /mol_type="other DNA"
               /db_xref="taxon:301574"
               /note="CMV 35S promoter"
               /note="CMV 35S promoter"
               /note="5' UTR from tobacco etch virus; translational
               enhancer"
               /note="N-terminal end of DsRed2 protein; part of fusion
               protein"
               /note="multiple cloning site; MCS"
misc_feature   2067..2277
terminator

```

CDS

/note="CamV 35S terminator"  
 complement(3486..4346)  
 /note="amp resistance"  
 /codon\_start=1  
 /product="Bla"  
 /protein\_id="AAV97912.1"  
 /db\_xref="GI:5653575"  
 /translation="MSIQHFRVALIPFPAAFCLEVPFAHPETLVKVKDAEDQLGARVGY  
 IEDLNSGRKILESRRPFRPMSTFKVLGCVLSRIDAOEQGLRRIRHNSQDLVE  
 YSPVTERKLTGGMTVRELCSAATMSDNTANLLLTIGRRELTAFLHNGDVTPL  
 DRWPELNEAI PNDERDTMPVMAATTLRKLTGELLTLASRQOLIIMMEDAKVAGPL  
 LRSALPAGMFIADKSGAGERSGRIIALAGDPKPSRIVIVYTTGSAITMDERRRQIA  
 EIGASLIRKM"

ORIGIN

Alignment Scores:

Pred. No.:	1.03e-110	Length:	4546
Score:	1186.00	Matches:	219
Percent Similarity:	98.22%	Conservative:	2
Best Local Similarity:	97.33%	Mismatches:	4
Query Match:	97.69%	Indels:	0
DB:	11	Gaps:	0

US-10-006-922A-12 (1-225) x AY818375 (1-4546)

1 MetArgSerSerLybAanValIleLygLuPhneMetArgpHeLybValArgMetGluGly 20  
 1333 ATGGCTCTCCGGAACGTCATCAACCGAGTTCAGCGCTTCAGGTGCGCATGGAGGAC 1382  
 21 ThrValaAngLybHleGluPhneGluIleGluGlyGluGlyArgProTyArgGluGly 40  
 1383 ACCGTGAACGGCCACGAGTTCAGATCGAGCGCGAGGGCGCGCCCTACAGAGGGC 1442  
 41 HisAenThrValLybLeuLybValThrLybGlyGlyProLeuProPhaIleTrpAspIle 60  
 1443 CACAACACCGTAGAGCTGAAGGTGACCAAGCGCGCGCCCTGCTGCTGCGGAGCATC 1502  
 61 LeuSerProGluPhneGluIleGlySerLybValThrValLybHisProAlaAspIlePro 80  
 1503 CTGTCCCTCCGAGTTCAGTACGCGCTCCAAAGGTACGTGAAGCACCCGCGACATCCCC 1562  
 81 AspTyLybLybLeuSerPheProGluGlyPheLybTyArgLybValMetAsnPhenGlu 100  
 1563 GACTACGAAGAGTCTCTCTCCCGAGGGCTTCAAGTGGAGCGGTGATGAACTTCAG 1622  
 101 AspGlyLybValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTy 120  
 1623 GACGGCGCGGTGGCGACCGGTGACCCAGGACTCTCTCCGTGACGAGCGGCTGCTCATC 1682  
 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGluIleTyThr 140  
 1683 AAGGTGAAGTTCATGGCGGTGAATCTCCCTCCGAGCGGCCCGGTATCAGAGAGAGCC 1742  
 141 MetGlyTyArgLybAlaSerThrGluArgLeuTyProArgAspGlyValLeuLybGlyGlu 160  
 1743 ATGGCTGGAGGCTCCACGAGCGCTGTACCCCGGACGCGGTCTGTAAGGGCGAG 1802  
 161 IleHisLybValAlaLeuLybLeuLybAspGlyGlyHisTyLeuValGluPhenLybSerIle 180  
 1803 ACCCAAGAGCGCTGAGCTGAAGGACGCGGCGCACTACGTGTGAGATTCAAGTCCATC 1862  
 181 TyrMetAlaLybValProValGluIleLeuProGlyTyTyTyValAspSerLybLeuAsp 200  
 1863 TACATGGCCAGAGAACCCGTGACGCTGCGGCTGCTACTACGTAGGAGCCAGCTGAGAC 1922  
 201 IleThrSerHisAsnGluAspTyThrIleValGluGluIleTyArgLybThrGluGlyArg 220  
 1923 ATCACTCCCAACAAGAGGACTACACATCGTGAAGCATGACGAGCGACCGAGGGCGGC 1982  
 221 HisHisLeuPheLeu 225  
 1983 CACCACCTGTTCTG 1997

RESULT 33

LOCUS DQ005468 4555 bp DNA circular SYN 07-MAY-2005

DEFINITION Cloning vector pSAT6A-DsRed2-N1, complete sequence.

ACCESSION DQ005468

VERSION DQ005468.1 GI:63002530

KEYWORDS

SOURCE

ORGANISM

Cloning vector pSAT6A-DsRed2-N1  
 Cloning vector pSAT6A-DsRed2-N1  
 other sequences; artificial sequences; vectors.

REFERENCE

1 (bases 1 to 4555)  
 Chung,S.-M. and Tzfira,T.  
 A versatile vector system for multiple gene expression in plants  
 Unpublished

JOURNAL

REFERENCE

2 (bases 1 to 4555)  
 Chung,S.-M., Yvas,S. and Tzfira,T.  
 Direct Submision  
 Submitted (12-APR-2005) Department of Biochemistry and Cell  
 Biology, State University of New York at Stony Brook, Stony Brook,  
 NY 11794, USA

FEATURES

source

Location/Qualifiers

1..4555  
 /organism="Cloning vector pSAT6A-DsRed2-N1"  
 /mol\_type="other DNA"  
 /db\_xref="taxon:326511"  
 441..762  
 /note="CamV 35S promoter"  
 768..1089  
 /note="CamV 35S promoter"  
 1190..1320  
 /note="translational enhancer 5'-UTR from tobacco etch  
 virus"

misc\_feature

1321..1379  
 /note="multiple cloning site; MCS"  
 1380..2057  
 /codon\_start=1  
 /product="DsRed2"  
 /protein\_id="AAV25372.1"  
 /db\_xref="GI:63002531"  
 /translation="MASSENVITPEFMKFKYRMETVNGHEFEIGEGEGRYEGHNTY  
 KLVTKGGLPFPAMDILSPQYGSKVYKHPADIPOYKULSPFEGFKMERKVMFEDG  
 GVALVTDSSLDGCFYKVKFIGNVPSDGPVWQKTMWEASTERLYPRDGLKGE  
 THKALKKDGSHYLVFEKSIYMAKKPVQLPGYVVDKLDITSHNBDYITVEQYERTE  
 GRHHLFLY"

2076..2286  
 /note="CamV 35S terminator"  
 complement(3495..4355)  
 /note="amp resistance"  
 /codon\_start=1  
 /product="Bla"  
 /protein\_id="AAV25373.1"  
 /db\_xref="GI:63002532"  
 /translation="MSIQHFRVALIPFPAAFCLEVPFAHPETLVKVKDAEDQLGARVGY  
 IEDLNSGRKILESRRPFRPMSTFKVLGCVLSRIDAOEQGLRRIRHNSQDLVE  
 YSPVTERKLTGGMTVRELCSAATMSDNTANLLLTIGRRELTAFLHNGDVTPL  
 DRWPELNEAI PNDERDTMPVMAATTLRKLTGELLTLASRQOLIIMMEDAKVAGPL  
 LRSALPAGMFIADKSGAGERSGRIIALADPDKPSRIVIVYTTGSAITMDERRRQIA  
 EIGASLIRKM"

ORIGIN

Alignment Scores:

Pred. No.:	1.03e-110	Length:	4555
Score:	1186.00	Matches:	219
Percent Similarity:	98.22%	Conservative:	2
Best Local Similarity:	97.33%	Mismatches:	4
Query Match:	97.69%	Indels:	0
DB:	11	Gaps:	0

US-10-006-922A-12 (1-225) x DQ005468 (1-4555)

1 MetArgSerSerLybAanValIleLygLuPhneMetArgpHeLybValArgMetGluGly 20  
 1380 ATGGCTCTCCGGAACGTCATCAACCGAGTTCAGCGCTTCAGGTGCGCATGGAGGAC 1439

QY 21 ThrValaIaenglyVhIsgluPhegluilegluglygluglygluglyAArgProTyrGluGly 40  
 Db 1440 ACCGGTGAACGGGCGACGAGTTCGAGATCGAGGGCGAGGGCGGCGCCCTTCAACGAGGAC 1499  
 QY 41 HIsaenThrVallyLeuLeuValThrlYsglyglYProLeuProPhaIaTriPaaplle 60  
 Db 1500 CACAACACCGGAGACTGAAAGGTGACCAAGGGGGGCCCCCTGGCCCTTGGGAGACATC 1559  
 QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrVallyVhIspRoIaAaspIlePro 80  
 Db 1560 CTGTCCCGCCCACTTCAGTACGCGCTCCCAAGGTGTAAGTGAACACCCCGCGACATCCCC 1619  
 QY 81 AapTyrLylyLeuSerPheProgluglyPhelygTTPgluAArgValMetAanPheglu 100  
 Db 1650 GACTCAAGAAAGCTGCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATTCGAG 1679  
 QY 101 AapGlyglYValValThrValThrlGlnAapSerSerLeuGlnAapGlyCyapPheIleTyr 120  
 Db 1680 GACGGCGGGCGGCGACCGGTGACCAAGACTCTCTCCCTGAGAGCGGCTGCTTCAATCAC 1739  
 QY 121 LyvVallyPheIleGlyValAanPheProSerAapGlyProValMetGlnLylySThr 140  
 Db 1740 AAGGTGAAGTTCATCGCGCTGAATTCCTCCCTCGAGCGCCCGCTGATGCAAGAAAGACC 1799  
 QY 141 MetGlyTTPgluAIsaerThrlGluArgLeuTyrProArgAapGlyValleuLyglYglu 160  
 Db 1800 ATGGGCTGGGAGGCGCTCCACCGAGGCGCTGTACCCCGCGAGCGGCGTGTAAAGGGCGAG 1859  
 QY 161 ILeHslyValleuLylyLeuLyAapGlyglYhIstYrLeuValGluPheLySerIle 180  
 Db 1860 ACCCAAGAGCGGCTGAACTGAAAGGACGAGCGGCGCACTACCTGTGAGATTCAAGTCCATC 1919  
 QY 181 TyrMetAlaIylyLyPheProValGlnLeuProglYrTyrTyrValAapSerLylyAap 200  
 Db 1920 TACATGGCCCAAGAGCGCGGTGAGCTGCCGCTACTTACTGAGACCCCAAGCTGAGAC 1979  
 QY 201 IleTrnSerHIsaenglyAanGlyAapTyrThrlIleValGluGlnTyrGluArgThrlGluArg 220  
 Db 1980 ATCACTCCCAACAGAGGACTTACATCGTGGAGCAATGAGGACCGAGGGGCGCC 2039  
 QY 221 HIsHsleuPheLeu 225  
 Db 2040 CACCACTGTCTCTG 2054  
 RESULT 34  
 AY818373 4570 bp DNA circular SYN 12-APR-2005  
 LOCUS AY818373  
 DEFINITION Cloning vector pSAT6-DsRed2-N1, complete sequence.  
 ACCESSION AY818373  
 VERSION AY818373.1 GI:56553569  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCES  
 AUTHORS  
 TITLE  
 JOURNAL  
 PUMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES  
 source  
 1..4570  
 /organism="Cloning vector pSAT6-DsRed2-N1"  
 /mol\_type="other DNA"  
 /db\_xref="taxon:301572"

promoter 441..762  
 /note="CaWY 35S promoter"  
 promoter 768..1090  
 /note="CaWY 35S promoter"  
 misc\_feature 1190..1320  
 /note="5' UTR from tobacco etch virus; translational enhancer"  
 misc\_feature 1323..1394  
 /note="multiple cloning site; MCS"  
 CDS 1395..2072  
 /codon\_start=1  
 /product="DsRed2"  
 /protein\_id="AAV97910.1"  
 /db\_xref="GI:56553571"  
 /translation="MASENVITPERMPKVMMEGVNCHEREIEGEGRGPEYGHNTV  
 KLYTKGGLPFPADILISPOYGSKYVHKPADIPYTKLSFPRGPKMERVMNEDG  
 GVATVTDSSLDGCFYKVFIGNPSPDVPVQKTMGEASTERYLPEDGLKGS  
 THKALKDGGHYLVEFKSIYMAKKPVQLPEYVYVDAKLDTSHNEDYTIYEYERTE  
 GRHLFL"  
 terminator 2091..2301  
 /note="CaWY 35S terminator"  
 complement(3510..4370)  
 /note="amp resistance"  
 /codon\_start=1  
 /product="bla"  
 /protein\_id="AAV97909.1"  
 /db\_xref="GI:56553570"  
 /translation="MSIGHFVALIPFPAFCLEVPFAHPETLVKKYDAEDQLGARVGY  
 IELDNLTKIESFPERFPMMSTFPTKLCGAVLSRIDAOEQVGRIRHYSQNDLVE  
 YSPTEKHLDTGMFRELCSAATWSDTANLLLTITGSKELTAPFANNGDHVPL  
 DRWEPELNEAI PNDERDPTMPVAMATLTLLTLLTASRQOLI DMWEADKVGPL  
 LRSALPACMPTIADSGAGBRSGRIIAGLSDGKRSRLVIVITTSQASQTMDBRRNQIA  
 EIGASLIKHW"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,04e-110 Length: 4570  
 Score: 1186.00 Matches: 219  
 Percent Similarity: 98.22% Conservative: 2  
 Best Local Similarity: 97.33% Mismatches: 4  
 Query Match: 97.69% Indels: 0  
 DB: 11 Gaps: 0  
 US-10-006-922A-12 (1-225) x AY818373 (1-4570)  
 QY 1 MetArgSerSerLyAanValIleLysgluPheMetArgPheLyValArgMetGluGly 20  
 Db 1395 ATGGCTCTCTCCGAGAGCTCATCAACGAGTTCATGCGGCTTCAAGGTGGCATGAGGGC 1454  
 QY 21 ThrValaIaenglyVhIsgluPhegluilegluglygluglygluglyAArgProTyrGluGly 40  
 Db 1455 ACCGTGAACGGGCGACGAGTTCGAGATCGAGGGCGAGGGCGGCGCCCTTACAGAGGGC 1514  
 QY 41 HIsaenThrVallyLeuLeuValThrlYsglyglYProLeuProPhaIaTriPaaplle 60  
 Db 1515 CACAACACCGTGAAGTGAAGGTGACCAAGGGGGGCCCCCTGGCCCTTCCCTGGGACATC 1574  
 QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrVallyVhIspRoIaAaspIlePro 80  
 Db 1575 CTGTCCCGCCCACTTCAGTACGCGCTCCCAAGGTGTAAGTGAACACCCCGCGACATCC 1634  
 QY 81 AapTyrLylyLeuSerPheProgluglyPhelygTTPgluAArgValMetAanPheglu 100  
 Db 1635 GACTCAAGAAAGCTGCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATTCGAG 1694  
 QY 101 AapGlyglYValValThrValThrlGlnAapSerSerLeuGlnAapGlyCyapPheIleTyr 120  
 Db 1695 GACGGCGGGCGGCGACCGGTGACCAAGACTCTCTCCCTGAGAGCGGCTGCTTCAATCAC 1754  
 QY 121 LyvVallyPheIleGlyValAanPheProSerAapGlyProValMetGlnLylySThr 140  
 Db 1755 AAGGTGAAGTTCATCGCGCTGAATTCCTCCCTCGAGCGGCGCGTGTATGCAAGAAAGACC 1814

Qy	141	MecGlyTrpGluAlaSerThrGluAlaXLeuYrProArGaSpGlyValAlleuYsGlyGlu	160
Db	1815	ATGGGCTGGAGAGGCTCCACCGAGCGCTGTATCCCGCGACGCGCTGTCTAAGAGGCGAG	1874
Qy	161	ILehIsylValAlaLeuYsLeuYsAspGlyGlyHsTrYLeuValGluPhelYsSerIle	180
Db	1875	ACCCACAAGGCGCTGAAGCTGAAGAGCGGGCGCACTACCTGGTGAAGTTCAAGTCCATC	1934
Qy	181	TyrMetAlaYbYsPProValGlnLeuProGlyYrYrYrYrValAspSerIysLeuAsp	200
Db	1935	TACATGGCCCAAGAGCCCGTGCACACTGCGCCGCTACTACTACGTGACGCGCAAGCTGGAC	1994
Qy	201	ILeHsSerHsAsnGluAspYrThrIleValGlnGlnTrYrGluYrGluThGluGlyAArg	220
Db	1995	ATCACCTTCCCAACAAGAGACTTACCATCTGTGAGACATACGACGCAAGAGGCGCGC	2054
Qy	221	HisHisLeuPheLeu	225
Db	2055	CACCACTCTTCTCTG	2069
RESULT 35			
LOCUS	AY640628/c	4829 bp	DNA
DEFINITION	sRNA vector pSUPER-hsYn-DsRed2N1-CyTB-AS, complete sequence.		
ACCESSION	AY640628		
VERSION	AY640628.1	GI:56119185	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
rep_origin			
promoter			
polya_signal			
intron			
gene			
CDS			
misc_feature			
promoter			

	promoter	misc_feature	promoter	rep_origin	gene	CDS
	to +53)"	2240..2455	/note="disabled human H1 RNA promoter"	2456..2587	/note="CytB-AS; partial human cytochrome B antisense sequence"	complement(2640..2658)
					/note="T3 promoter"	3026..3593
					/note="pUC Ori"	/note="ampicillin resistance"
					/gene="Amp"	complement(3841..4701)
					/gene="Amp"	complement(3841..4701)
					/note="Amp"	
					/note="ampicillin resistance"	
					/codon_start=1	
					/transl_table=11	
					/product="beta-lactamase"	
					/protein_id="AAV73959.1"	
					/db_xref="GI:56119187"	
					/translation="MSIQHRAVLIPFAFLCPVAHPETLVKQDAEQLGARVY	
					IELDLSNGIKLESFPEERFPMWSTFKVLGCVLSRIDAGQQLERLRHYSQNDLVE	
					YSSVETKHLLTDGWTVELSCSAITMSDNPANILFTTIIGPKKLTFLHMGSHVRL	
					DREPELNEAIPIDESDITMPVAMATTLRLLTGLLTLASROQLIDMEADKYAGPL	
					LRBALPAGWFIADKSGAGRGSRGIIAALGDPKSPRIIVITYTTSQATMDENRQIA	
					EIGASLIKRW"	
	ORIGIN					
	Alignment Scores:					
	Pred. No.:	1,11e-110	Length:	4829		
	Score:	1186.00	Matches:	219		
	Percent Similarity:	98.22%	Conservative:	2		
	Best Local Similarity:	97.33%	Mismatches:	4		
	Query Match:	97.69%	Indels:	0		
	DB:	11	Gaps:	0		
	US-10-006-922A-12 (1-225) x AY640628 (1-4829)					
QY	1 MetArgSerSerIysAsnValIleYsGluPheMetArgPheIysValArgMetGluGly	20				
Db	1722 ATGAGCTCTCCGAGAACGTCATCAACCGAGTTTCATCGCTTCAAGGCGCGCATGAGGGC	1663				
QY	21 ThrValAsnGlyIleGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40				
Db	1662 ACCGTGAACGCCACGAGTTCGAGATCGAGGGCGAGGGCCAGGGCCGCTTACGAGGGC	1603				
QY	41 HisAsnThrValIysLeuIysValIleThrIysGlyGlyProLeuProPheIleTyrPheIle	60				
Db	1602 CACAAACCGTGAAGCTGAAGGTGACCAAGGGGGCCCTTCCGCTTGGCTGGGACATC	1543				
QY	61 LeuSerProGluPheGluIleTyrGlySerIysValIleTyrValIleHisProIleAspIlePro	80				
Db	1542 CTGTCCCCCAGTTTCCAGTACGGCTCCAGAGTGTACGTGAAGACCCCGCGCATCTCCC	1483				
QY	81 AspTyrIleIysLeuSerPheProGluGlyPheIleTyrGluArgValMetAsnPheGlu	100				
Db	1482 GACTACAAAGAACGTCTCTTCCCGAGGGCTTCAAGTGGGAGGCGCGATGAATCTTCAG	1423				
QY	101 AsnGlyGlyValIleThrValIleThrGluAsnSerSerLeuGluAsnGlyCysPheIleTyr	120				
Db	1422 GACGGCGCGCTGGCGAACCGTGAACCGAGTCTCTTCCGTGAGAGCGGCTGCTCATTTAC	1363				
QY	121 IysValIysPheIleGlyValIleAsnPheProSerAspGlyProValMetGluIysThr	140				
Db	1362 AAGTGAAGTTTCAATCGCGGTGAATCTTCCCTCCGAGGCGCCCGTGAAGCAGGAAGAC	1303				
QY	141 MetGlyTyrGluIleAsnThrGluArgLeuTyrProArgAspGlyValIleIysGlyGlu	160				
Db	1302 ATGGGCTGGAGGCTCTTCCACCGAGCGCTGTATCCCCCGCAGCGCGCTGGAAGGGCGAG	1243				
QY	161 IleHisIysValIleIysLeuIysAsnArgIysGlyHisTyrLeuValGluPheIysSerIle	180				
Db	1242 ACCCAAGGCGCTGAAGCTGAAGAGACGGGGCGACATCACTGTGTGAGTTCAAGTTCATC	1183				

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAlaSerLysLeuAsp 200  
Db 1182 TACATGGCCAAAGAACCCGTCAGCTGCGCCGCTACTACTAGTGGACGCCAACGCTGGAC 1123  
Qy 201 TleThSerHisAenGluAspTyrThrIleValGluGlnTyrGluLysGlnThrGluGluArg 220  
Db 1122 ATCACTCCCAACAGAGACTACACATCTGTGAGAGACTAGACGACGACGAGGCGCG 1063  
Qy 221 HisHisLeuPheLeu 225  
Db 1062 CACCACCTGTTCTCTG 1048  
RESULT 36  
LOCUS AY640625 4906 bp DNA circular SYN 01-JUL-2005  
DEFINITION S1RNA vector pSUPER-CMV-DsRed2N1-Cytrb-AS, complete sequence.  
ACCESSION AY640625  
VERSION AY640625.1 GI:56119176  
KEYWORDS  
SOURCE s1RNA vector pSUPER-CMV-DsRed2N1-Cytrb-AS  
ORGANISM s1RNA vector pSUPER-CMV-DsRed2N1-Cytrb-AS  
REFERENCE 1 (bases 1 to 4906)  
AUTHORS Michel,U., Malik,I., Ebert,S., Bahr,M. and Kugler,S.  
TITLE Long-term in vivo and in vitro AAV-2-mediated RNA interference in rat retinal ganglion cells and cultured primary neurons  
JOURNAL Biochem. Biophys. Res. Commun. 326 (2), 307-312 (2005)  
PUBMED 1582578  
REFERENCE 2 (bases 1 to 4906)  
AUTHORS Michel,U. and Kugler,S.  
TITLE Direct Substitution  
JOURNAL Submitted (27-MAY-2004) Neurology, University of Goettingen, Walweg 33, Goettingen 37073, Germany  
FEATURES  
source 1..4906  
/organism="s1RNA vector pSUPER-CMV-DsRed2N1-Cytrb-AS"  
/mol\_type="other DNA"  
/db\_xref="taxon:297627"  
135..441  
/note="PI-Or1"  
rep\_origin 625..645  
/note="T7 promoter"  
promoter complement(739..899)  
/note="SV40-PA; SV40 polyadenylation sequence"  
polyA\_signal complement(899..1039)  
/note="SV40 chimeric Intron derived from pCI-NEO"  
intron complement(1045..11722)  
/gene="DsRed2"  
gene /note="derived from Clontech's pDsRed2-N1"  
complement(1045..11722)  
CDS  
/gene="DsRed2"  
/codon\_start=1  
/transl\_table=11  
/product="DsRed2"  
/protein\_id="AAV73952.1"  
/db\_xref="GI:56119177"  
/translation="MASSENVITERPFRFKYRMGGTYNGHFRTEGSGRPRYGHNTV  
KIKVTGSPPLPFAWDILSPQFGSKYVYAHPADIDYKLSFPEPFKRYNNPFDG  
GVATVQDPSLDGCPRIYKVKFIGNVPDGPWQKSGWEXSTRLPRDVLKGE  
THALALIKDGGHYLVEFKSIYMAKKPVOLPGYVYDAKIDITSHNEDYITVEQYKTE  
GRHHLPL"  
complement(1741..1760)  
misc\_feature /note="MCS"; multiple cloning site"  
promoter complement(1771..2293)  
/note="MCMV promoter; mcmv promoter (-491 to +36) "  
misc\_feature 2294..2309  
/note="MCS"; multiple cloning site"  
promoter 2317..2532  
/note="disabled human H1 RNA promoter"  
misc\_feature 2533..2664  
/note="Cytrb-AS; partial human cytochrome B antisense  
sequence"

promoter complement(2717..2735)  
/note="T3 promoter"  
rep\_origin 3103..3770  
/note="PUC Ori"  
gene complement(3918..4778)  
/gene="Amp"  
CDS complement(3918..4778)  
/note="Amp"  
/gene="ampicillin resistance"  
/codon\_start=1  
/transl\_table=11  
/product="beta-lactamase"  
/protein\_id="AAV73953.1"  
/db\_xref="GI:56119178"  
/translation="WSIOHFRVALIPPPAFCLEVPFAHPETLVYVKAAEDQGARVGY  
IELDINSKILIESFRPERFPWMTFKLLCGAVALSRIDAQEQUGRIHYSQNDLVE  
YSPVKEHLTGDMTRELCSNAITMSDNTANLLITYGKRELTAFLANNGDHVTPL  
DRWPEINBEAI PNDERDPTMVPAMATTLTKLTGELTLTASRQQLIDWMEADKVAGPL  
LRSLPAGMFIADSGNGERSRGI IALGPDKPSRLVIYTTGSGATMDERNRQIA  
ETGASLIRKM"

ORIGIN  
Alignment Scores:  
Pred. No.: 1,136-110 Length: 4906  
Score: 1186.00 Matches: 219.  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
Gaps: 0  
DB: 11  
US-10-006-922a-12 (1-225) x AY640625 (1-4906)  
Qy 1 MetArgSerSerLysAsnValIleLeuGluPheMetArgPheLeuValArgMetGluGly 20  
Db 1722 ATGGCTCCTCCCGAAGACGTCACTCAACGATTCATGCGCTTCACAGTGGAGGCG 1663  
Qy 21 ThrValaenGlyHisGluPheGluIleGluGluGluGluGluGluGluGluGluGluGlu 40  
Db 1662 ACCGGAACGGCCAGAGTTCAGATTCAGAGGCGAGGCGAGGCGAGGCGAGGCGAGGCG 1603  
Qy 41 HisAenThrValaLysLeuLeuValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 1602 CACCAACCGCTGAAGCTGAAGTGAACAGGCGGCGCCCTGCTGCGCGGAGCATC 1543  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValaHisProAlaAspIlePro 80  
Db 1542 CTGTCCCGCCCAATCCAGTACGAGTCCAAAGGTGTAGTGAACACCCCGCACATCCCC 1483  
Qy 81 AspTyrLysLysLeuSerPheProGluGluPheLysTrpGluArgValMetAsnPheGlu 100  
Db 1482 GACTACMAAGACTGTCTTCCCGAGGCGCTTCAAGTGGAGCGGCGTGAATGAACTTCGAG 1423  
Qy 101 AspGlyGlyValaValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 1422 GACGGCGCGCTGCGACCGTGAACCGACCTCTCTCCCTCAAGAGCGCTGCTTCACTAC 1363  
Qy 121 LysValaLysPheIleGlyValaAsnPheProSerAspGlyProValMetGlnLysTyr 140  
Db 1362 AAGGTAAATTATGCGCGTGAACCTCCCTCGACGCGCCCGTGAATGCAAGAACGACC 1303  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValaLeuLysGlyGlu 160  
Db 1302 ATGGGCTGGAGGCGCTCCACGAGCGCGCTGTACCCCGGAGCGCGCTGCTGAAGGCGAG 1243  
Qy 161 HisHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 1242 ACCCAAGAGCCCTGAAGCTGAAGAGCGGCGGCACTACTCGGTGAAGTCAAGTCATC 1183  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAlaSerLysLeuAsp 200  
Db 1182 TACATGGCCAAAGAACCCGTCAGCTGCGCCGCTACTACTAGTGGACGCCAACGCTGGAC 1123  
Qy 201 TleThSerHisAenGluAspTyrThrIleValGluGlnTyrGluLysGlnThrGluGluArg 220

```

Db      1122 ATCACTCCCAAGAGAGCTACACCATCTGTGAGCATGAGACGACCGAGGCGCC 1063
Oy      221 H1sh1sleupheleu 225
Db      1062 CACCACTGTTCTTG 1048

RESULT 37
LOCUS   AJ851284                      5311 bp      DNA      circular SYN 12-NOV-2004
DEFINITION Cloning vector pRU1104 dsRed2 gene for Ds Red fluorescent protein
ACCESSION AJ851284
VERSION   AJ851284.1 GI:55724874
KEYWORDS  Ds Red fluorescent protein 2; dsRed2 gene.
SOURCE    Cloning vector pRU1104
ORGANISM  Cloning vector pRU1104
           other sequences; artificial sequences; vectors.
REFERENCE 1
AUTHORS   Karunakaran, R. and Poole, P.S.
TITLE     High throughput promoter probe vectors
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 5311)
AUTHORS   Poole, P.S.
TITLE     Direct Submision
JOURNAL   Submitted (21-OCT-2004) Poole P.S., School of AMS, University of
           Reading, Whiteknights, Reading, RG6 6AU, UNITED KINGDOM
FEATURES
     source             1..5311
                        /organism="Cloning vector pRU1104"
                        /mol_type="other DNA"
                        /db_xref="taxon:299179"
                        /lab_host="Escherichia coli"
                        /gene="dsRed2"
                        /protein_id="CAH64889.1"
                        /codon_start=1
                        /product="Ds Red fluorescent protein 2"
                        /transl_table=11
                        /protein_id="CAH64889.1"
                        /db_xref="GI:55724875"
                        /translation="MASSENVITEFMRFKVMEGTVNGHFEIEGEGGRPYEGHNTV
                        KLKTKGGLPFPAMDILSPROYGSKVVKPADIPDKXLSFPGFKMEVMPEDG
                        GVATVDDSLDGCGRYKVKFKIGNPNPDKWOKMGEASRERLYPDDVILKGR
                        THKALKDKDGGHYLVFKSIYMAKKPVLPQYIVDAKLDITSHNEDTIVYQYRTE
                        GRHHLPL"

ORIGIN
Alignment Scores:
Pred. No.:      1,256-110      Length:      5311
Score:          1186.00      Matches:      219
Percent Simlarity: 98.22%      Conservative: 2
Best Local Simlarity: 97.33%      Mismatches: 4
Query Match:    97.69%      Indels:      0
DB:             11      Gaps:      0

US-10-006-922a-12 (1-225) x AJ851284 (1-5311)

Oy      1 MetAArgSerSerLyAsnValIleLySGIuphMetAArgPheLyValArgMetGIuLy 20
Db      326 ATGGCTCTCCGAGAACGTATCAACGAGTTCAATGCGCTTCAAGTGCATGAGGCG 385
Oy      21 ThrValaAngLyHieGIupheGIuIleGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 40
Db      386 ACCGGAACGGGCAACGAGATTCAGAGTCAGAGGCGGAGGCGGCGCCCTACAGAGGC 445
Oy      41 HisAsnThrValLyLeuLyValThrLySGIyGIyProLeuProPhaIaTpaAspIle 60
Db      446 CACAACACCGGAGAGCTGAGGTGACCAAGGCGGCGCCCTTGCTTGCGGAGATC 505
Oy      61 LeuSerProGIuPhieGIuTyrGIySerLyAsnValTyrValLyShAProAlaAspIlePro 80

```

```

Db      506 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGACACCCCGCAGATCCCC 565
Oy      81 AspTyrLySlySleuSerPheProGIuGIyPheLySerTTPGuaGvAlMetAsnPhieIu 100
Db      566 GACTTACAAGAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGGTGATGAACCTTCAG 625
Oy      101 AspGIyGIyValValThrValThrGIuAspSerSerLeuGIuAspGIyCySPhaIleTyr 120
Db      626 GACCGCGCGCGTGGAGACCGTGAACCAAGATCTCTCCCTGACGAGACGGCTTCACTAC 685
Oy      121 LyValLySPhaIleGIyValaAsnPheProSerAspGIyProValMetGIuLyLySerThr 140
Db      686 AAGGTGAAGTTCAATCGCGCTGTAATCTCCCTCCGACGCGCCCGTGAATGCAAGAAAGACC 745
Oy      141 MetGIyTTPGuaIaSerThrGIuArgLeuLyTyrProArgAspGIyValIleuLyGIyGIu 160
Db      746 ATGGCTGGAGGCGCTCCACGAGCGCTGTATCCCGCGACGCGCGTGTCAAGAGGCGAG 805
Oy      161 IleHisLyAlaLeuLySleuLyAspGIyGIyHisTyrLeuValGIuPhaLySerIle 180
Db      806 ACCCAAGAGCGCTCGAAGCTGAAGAGACGCGCGCACCTACCTGTGAGTTCAAGTCCATC 865
Oy      181 TyrMetAlaLySlySProValGIuLeuProGIyTyrTyrTyrValaAspSerLySleuAsp 200
Db      866 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGCTACTACTACGTGAGACGCCAAGCTGGAC 925
Oy      201 IleThrSerHisAsnGIuAspTyrThrIleValGIuGIuTyrGIuArgTThGIuGIyArg 220
Db      926 ATCACTCCCAAGAGAGACTACACATCGTGAAGACGTACGAGCGCACCGAGGCGCGC 985
Oy      221 H1sh1sleupheleu 225
Db      986 CACCACTGTTCTTG 1000

RESULT 38
AY613997                      6423 bp      DNA      circular SYN 26-JUN-2004
LOCUS   AY613997
DEFINITION Cloning vector pSRalaphaneOR, complete sequence.
ACCESSION AY613997
VERSION   AY613997.1 GI:48995616
KEYWORDS
ORIGIN   Cloning vector pSRalaphaneOR
SOURCE   Cloning vector pSRalaphaneOR
          other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6423)
AUTHORS   Allard, V., de Leseleuc, L. and Denis, F.
TITLE     A family of mammalian expression vectors with different selection
          markers
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 6423)
AUTHORS   Allard, V., de Leseleuc, L. and Denis, F.
TITLE     Direct Submision
JOURNAL   Submitted (01-MAY-2004) INRS-Institut Armand-Frappier, 531
          Boulevard des Prairies, Laval, Quebec H7V 1B7, Canada
FEATURES
     source             1..6423
                        /organism="Cloning vector pSRalaphaneOR"
                        /mol_type="other DNA"
                        /db_xref="taxon:279783"
                        /note="ECMV internal ribosomal entry site"
                        /note="ECMV internal ribosomal entry site"
                        /note="ECMV internal ribosomal entry site"
                        /gene="dsRED2"
                        /gene="dsRED2"
                        /gene="dsRED2"
                        /codon_start=1
                        /transl_table=11
                        /product="red fluorescent protein"
                        /protein_id="AAT48428.1"
                        /db_xref="GI:48995617"
                        /translation="MASSENVITEFMRFKVMEGTVNGHFEIEGEGGRPYEGHNTV

```



```

misc_feature
     2134..2266
     /note="contains SV40 polyA signal"
     complement(2294..3088)
gene
     /gene="aphII"
     complement(2294..3088)
     /gene="aphII"
     /codon_start=1
     /transl_table=11
     /product="neomycin phosphotransferase"
     /protein_id="AA148429.1"
     /db_xref="GI:48995618"
     /translation="MIEDGLHAGSPAAWVERLPGYDWAQOTIGSDAVERLSAQR
PVLPRKTLTSLANLQDPAARLSTLPAATCPFHQAKHIERAKTMEGLVDQDDE
LSHIAPEAKSIVADAKRLHTLDPATCPFHQAKHIERAKTMEGLVDQDDE
LHOGIAPAPLPAELKAPMDGDLVVTGDAFLPNIWENFSGFIDGRLGVADRY
ODIALATRDIAELGEMADRLVYGIAPDSQRIARVRLDEPF"
     complement(3456..3783)
     /note="SV40 early"
promoter
     /note="SV40 early"
     3914..4774
     /gene="b1a"
     3914..4774
     /gene="b1a"
     /codon_start=1
     /transl_table=11
     /product="ampicillin resistance protein"
     /protein_id="AA148430.1"
     /db_xref="GI:48995619"
     /translation="MSIQHFRVALIPFPAFCLPVFAHPETLVKVKDAEDQLGARVY
TELDNSGKILSPERFPMWSTPKVLGCAVSRVDAQGEQGRIRHYSDNLTVE
YSPVETKHLTDKMTRELCSAATWSNDTAAALITLTGGPELTAAPLWMDKVAIPL
DMEPELNAIPNDERDTTPAAAITTLAKLTIGELITLASQQLIDMWEDKVAIGL
LRSALPAGWFIADKSGAGERSGRIIALLGPDKSRIVVITTSQATMDERNQIQA
ETGASLIKIM"
     4961..5535
     /note="colE1"
     5786..6408
     /note="Shalpa"
rep_origin
promoter
ORIGIN
Alignment Scores:
pred. No.:      1,57e-110      Length:      6423
Score:          1186.00        Matches:      219
Percent Similarity: 98.228      Conservative: 2
Best Local Similarity: 97.338      Mismatches: 4
Query Match:    97.694         Indels:      0
DB:             11            Gaps:        0
US-10-006-922A-12 (1-225) x AY613997 (1-6423)
QY      1 MetArgSerSerIySaAnValIlelyGluPhmeLArgPheLyVaIArgMeCglUgly 20
      1431 ATGGCCTCTCCCGAAGACGTCAACCGAGTTCATCGCTTCAAGGTGCGCATGAGGCG 1490
QY      21 ThrValArgGlyHtIsgLupheGluIleGluGlyGluGlyArgProTyArgUgly 40
      1491 ACCGTAAACGGCCACGAGTTCGAGATCGAGGCGGAGGCGGCGCCCTTACGAGGCG 1550
QY      41 HLeaenThrValIySeuLyVaIThrLyGlyGlyProLeuProPhealATPaapIle 60
      1551 CACAAACCTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCGCTTGGGACATC 1610
QY      61 LeuSerProGlnPheGlnTyrgIySerLyVaIlyrValIyshiSProAlaapIlePro 80
      1611 CTGTCCCGCCAGTTCACGAGCTCCCAAGGTGACGTGAAGCACCCCCGGACATCCCC 1670
QY      81 AspTyIlyAllyLeuSerPheProGluGlyPheLyfTrgIuArgValMeTaaPheGlu 100
      1671 GACTACAAAGAGCTGTCTTCCCGGAGGCTTCAAGGTGAGCGCGGTGATGAACCTTCAG 1730

```

```

QY      101 AspGlyGlyValValThrValThrGluAspSerLeuGlnAapGlyCyPhePheIleTy 120
      1731 GACGCGCGCGGCGGACCGTACCAAGACCTCTCCCTGCGAGACGCTGCTTATCTAC 1790
QY      121 LySeValIyPheIleGlyVaIaSnPheProSeArApGlyProValMeGlnLyLyThr 140
      1791 AAGGTGAAGTTCATCGGCTGAACCTTCCCTCCGACGCGCCCGGTGATCAAGAAAGAC 1850
QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyProAlaGaaPglyValLeuLyGlyGlu 160
      1851 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGGCGTCTGAAGGCGCAG 1910
QY      161 HLeHlySaIaLeuLySeuLySaPglYgLyHstYrLeuValGluPheLySerIle 180
      1911 ACCCAAGGCGCCGTGAAGCTGAAGGCGGCGGCACTACCTGTGGAGATTCAAGTCCATC 1970
QY      181 TyrMeAlaLyLyPProValGlnLeuProGlyTyTyTyTyValaIaSPSerLySeuAap 200
      1971 TCAATGCCAAGAACCCGCTGACGCTGCCGCTACTACTACGTGACGCCAAGCTGGAC 2030
QY      201 HLeHrSerHLeaenGluAapTyThrIleValGluGlnTyrgIuArgThrGluGlyArg 220
      2031 ATCACTCCCAACGAAGACATCACTGTGAGCAGTACGAGCGACCGAGGCGCGCG 2090
QY      221 HSHISLeuPheLeu 225
      2091 CACCACTGTCTCTG 2105
DB
RESULTS 39
AY640634/c 6990 bp DNA circular SYN 01-JUL-2005
LOCUS AY640634
DEFINITION AY640634
ACCESSION AY640634
VERSION AY640634.1 GI:56119203
KEYWORDS
SOURCE
ORGANISM
sirNA vector pAAV9(5)-hSyn-DsRed2N1-BGFP-sirNA
sirNA vector pAAV9(5)-hSyn-DsRed2N1-BGFP-sirNA, complete sequence.
Other sequences; artificial sequences; vectors.
REFERENCE
1 (bases 1 to 6990)
Michel,U., Malik,I., Ebert,S., Bahr,M. and Kugler,S.
Long-term in vivo and in vitro AAV-2-mediated RNA interference in
rat retinal ganglion cells and cultured primary neurons
Biochem. Biophys. Res. Commun. 326 (2), 307-312 (2005)
JOURNAL
PUBMED
15582578
REFERENCE
2 (bases 1 to 6990)
Michel,U., Malik,I. and Kuegler,S.
Direct Submision
Submitted (01-JUN-2004) Neurology, University of Goettingen,
Waldweg 33, Goettingen 37073, Germany
FEATURES
Location/Qualifiers
1..6990
/organism="sirNA vector pAAV9(5)-hSyn-DsRed2N1-BGFP-sirNA"
/mol_type="other DNA"
/db_xref="taxon:297636"
complement(38..193)
/note="SV40-pA ; SV40 polyadenylation sequence"
complement(198..341)
/note="SV40 chimeric intron derived from pCI-NEO"
complement(348..1025)
/gene="DsRed2"
/note="derived from Clontech's pDsRed2-N1"
complement(348..1025)
/gene="DsRed2"
/codon_start=1
/transl_table=11
/product="DsRed2"
/protein_id="AAV73970.1"
/db_xref="GI:56119204"
/translation="MASSENVITTEPRFVRMEGTYNHGEFIEGEGSRPYGCHNTY
KLVTKGGLPFPAMDILSPQFYGSKVYVHPADIIPDYKLSPPBGFKEWVWNPEDG
GVAIVTQDSLDGCFIYKVFIGNVFPDGDVMOKKTGWSASTERLIPRQGVLDGE
THKALKIKDGGHYLVFKSIYMAKPKVQLPGYYVADAKLIDTSHNEDYITVEQYERTE
GRHHLPL"

```



```

promoter      /note="WCS; multiple cloning site"
               complement(1205..1674)
               /note="HSYN-promoter; human synapsin 1 gene promoter (-422
               to +53)"
promoter      1686..1901
               /note="disabled human H1 RNA promoter"
misc_feature  1902..2033
               /note="Cytb-AS; partial human cytochrome B antisense
               sequence"
gene          complement(2038..4301)
               /gene="9(5)"
               /note="partial sequence from the non-coding porcine RNA UM
               9(5)"
repeat_region 4321..4461
               /note="right ITR"
               /rpt_family="Inverted"
gene          5378..6238
               /gene="Amp"
               /gene="Amp"
               /gene="Amp"
               /note="ampicillin resistance"
               /codon_start=1
               /transl_table=11
               /product="beta-lactamase"
               /protein_id="AAV73968.1"
               /db_xref="GI:56119201"
               /db_xref="GI:56119201"
               /translation="MSIOHFRVALIPPPAFLPVPAHPELVKVKDAEDQLGARVGY
               /ELIDNSKILLESFRRPFPMWSFTFKVLGCAVVRIDAGEOGLRHHSONLVK
               VSPVTEKRLTDTMTRELCASAIITMSDNTANLLLTGGSPETLPAFLHMGDHTLV
               DRWPELNEALPNDERDITWPAVMATTKRLKLTGSLITVLASQQLIDMWEADKVAQL
               LRSALPACMFADSKAGERSRGIIALGPDGKPSRLVITTSQATMDERNRQIA
               EIGSLAIKHW"
ORIGIN
Alignment Scores:
Score:         1,766-110      Length:      7058
Percent Similarity: 1186.00    Matches:     219
Best Local Similarity: 98.22%   Conservative: 2
Query Match:    97.33%         Mismatches:  4
DB:             11            Indels:       0
               Gaps:         0
US-10-006-922A-12 (1-225) x AY640633 (1-7058)
QY      1  McArGSeSerLySaenValIleLySgIuPhMeArGpHeLyVaIaRgMeGluGly 20
DB      1168 ATGGCCTCTCCGAGAACGTCATCACCGAGTTCATCGCTTCAAGGTGGCGATGAGGGC 1109
QY      21 ThrValaSnGlyVHAgIuPhneGluIleGluGlyGluGlyValrGpProTyrgIuGly 40
DB      1108 ACCGTAAAGGCGACAGAGTTCAGATCGAGGGGAGGGGCGAGGGCGCCCTACGAGGCG 1049
QY      41 HlaenThrVallyLeuLyVaIThrLyGlyGlyProLeuProPhelaIatPaaPile 60
DB      1048 CACACACACCGTGAAGCTGAAGGACCAAGGGGGCGCCCTCGCTCGGAGCATC 989
QY      61 LeuSerProGILPhneGIntYrgIySerLyVaIlyVAllybHIsPaaIaAepIlePro 80
DB      988 CTGTCCCCCAAGTTCAGTACGAGCTCCAGGTACGTAAAGCAACCCCGCCGACATCCCC 929
QY      81 AApTyLyLyLeuSerPheProGluGlyPheLySgIuRgVaImeAenPheGlu 100
DB      928 GACTACAAAGAGCTGCTCTCCCGAGGGCTTCAAGTGGAGCGCTGATGAATCTCGAG 869
QY      101 AApGlyLyVaIValThrValThrGlnaPseSerLeuGlnaAepGlyCybPheIleTyR 120
DB      868 GACGGGGCGGTGGCGACCGTGAACCGAGACTCTCCCTCGAGAGCGAGCTGCTTCATCTAC 809
QY      121 LybVallybPheIleGlyValIaenPheProSerAapGlyProValMeGIntLybSThr 140
DB      808 AAGGTGAAGTTCAATCGGCGTGAACCTTCCCTCCGACGCGCCCTGATGCAAGAAAGACC 749
QY      141 MetGlyTrpGluIaAseThrGluArgLeuTyRProArgAapGlyValleuLySgIyGlu 160

```

```

DB      748 ATGGGCTGGAGGCTCTCACCGAGCGCTGTACACCCCGCAGCGCGTGTGAAGGCGAG 689
QY      161 IleHisLySaIaLeuLySleuLyAapGlyGlyHIsTyRleuVaIgluPhneLySerIle 180
DB      688 ACCCAAGAGCCCTGAAGCTGAAGAGACGCGGCCCACTACCTGGTGGAGTTCAAGTTCATC 629
QY      181 TyrMetAlaLybLybProValGlnLeuProGlyTyTyTyTyTyValIaApsSerLyLeuAap 200
DB      628 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCTACTACTACGTGAGACGCCAAGCTGGAC 569
QY      201 IleThrSerHisAenGluAapTyRThrIleValGluGIntYrgIuArgThrGluGlyArg 220
DB      568 ATCACTTCCCAACAGAGACTACATCCTGTGAGACGTACGAGCGACGAGGCGCCGC 509
QY      221 HisHisLeuPheLeu 225
DB      508 CACCACTGTCTCTG 494
RESULT 41
AY640630/C 7147 bp DNA circular SYN 01-JUL-2005
LOCUS      AY640630
DEFINITION S1RNA vector PAAV9 (5)-CMV-DsRed2N1-CytB-AS-ohneNot, complete
sequence.
ACCESSION AY640630
VERSION    AY640630.1 GI:56119191
KEYWORDS   S1RNA vector PAAV9 (5)-CMV-DsRed2N1-CytB-AS-ohneNot
SOURCE     other sequences; artificial sequences; vectors.
ORGANISM   1 (bases 1 to 7147)
REFERENCE  1 (bases 1 to 7147)
AUTHORS   Michel,U., Malik,I., Ebert,S., Bahr,M. and Kugler,S.
TITLE     Long-term in vivo and in vitro PAV-2-mediated RNA interference in
          rat retinal ganglion cells and cultured primary neurons
          Biochem. Biophys. Res. Commun. 326 (2), 307-312 (2005)
JOURNAL    PUBMED 15582578
REFERENCE  2 (bases 1 to 7147)
AUTHORS   Michel,U., Malik,I. and Kuegler,S.
TITLE     Direct Submission
          Submitted (01-JUN-2004) Neurology, University of Goettingen,
          Waldweg 33, Goettingen 37073, Germany
JOURNAL    FEATURES
          source
          1..7147
          /organism="S1RNA vector
          PAAV9 (5)-CMV-DsRed2N1-CytB-AS-ohneNot"
          /mol_type="other DNA"
          /db_xref="taxon:297632"
          1..141
          /note="left ITR"
          /rpt_family="Inverted"
          /note="SV40-pA; SV40 polyadenylation sequence"
          complement(181..338)
          complement(341..481)
          /note="SV40 chimeric intron derived from pCI-NEO"
          complement(491..1168)
          /gene="DsRed2"
          complement(491..1168)
          /note="DsRed2"
          /note="derived from Clontech's pDsRed2-N1"
          /codon_start=1
          /transl_table=11
          /product="DsRed2"
          /protein_id="AAV73963.1"
          /db_xref="GI:56119193"
          /translation="MASSENVITTEFMREKVRMEGTVNGHEPFIIEBGBGRPYEGHNTV
          KLVKRGKPLPPAMPILIDPQFGSKVYVHPADLPDYKLSFPFGFKRVMNVEDG
          GYATYQDSLDDGGCEIKYKTKRGIVNPSDGPVMQKTMGWASTBRLYPRGGVYKGE
          THKALKDGGHIVFESKITPAKRPVQLPGITYDADLDITSHNEDYIVQYERTE
          GRHHLFL"
          complement(1187..1206)
          /note="MCS; multiple cloning site"
          complement(1217..1739)
          /note="MCMV promoter; mcmv promoter (-491 to +36)"
repeat_region
polyA_signal
intron
gene
CDS
misc_feature
promoter

```

```
promoter      1775..1990
              /note="disabled human H1 RNA promoter"
misc_feature  1991..2122
              /note="Cytb-A5; partial human cytochrome B antisense
              sequence"
gene          2127..4390
              /note="partial sequence from the non-coding porcine RNA UM
              9 (5)"
repeat_region 4410..4550
              /note="right ITR"
              /rpt_family="inverted"
gene          5467..6327
              /gene="Amp"
              /length=860
              /note="Amp"
              /note="ampicillin resistance"
              /codon_start=1
              /product="beta-lactamase"
              /protein_id="AAV73962.1"
              /db_xref="GI:56119192"
              /translation="MSIQHFRVALIPPPAFCLPVFAHPETLVKVKQADQLGKRVY
              IELDINSKIGLSPERFPMWSTFKYLCAVLSDAQOQLGRIRIHYSQNDLVE
              YSPVTEKILDTGMIVRELCSAAITWSDNTANLLITIGPKRELTAFLHMGDHYRL
              DRMEPELNEAI PNDEKDTTPVMAATLTKLITGELTLASRQOLIDMEADKVAQPL
              LRSLLPACMPALADKSGAGERSRGITIALGPDGKSRIVIVYTTGSAQTMDERRRQILA
              ETGASLIRKM"
```

## ORIGIN

```
Alignment Scores:
Pred. No.:      1.79e-110      Length:      7147
Score:          1186.00      Matches:      219
Percent Similarity: 98.22%      Conservative: 2
Best Local Similarity: 97.33%      Mismatches: 4
Query Match:    97.69%      Indels:      0
DB:             11      Gaps:      0
```

US-10-006-922a-12 (1-225) x AY640630 (1-7147)

```
QY      1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20
DB      1168 ATGGCTCTCCGGAACAGTCATCAACGAGTTCATGCTCAAGGTGCGCATGAGAGGC 1109
QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB      1108 ACCGTGAACGGCCAGATTCAGATCGAGGCGAGGCGCGCCCTACGAGGCGC 1049
QY      41 HisAsnThrValLyLeuLyValThrLyGlyGlyProLeuProPheAlaTyrAspIle 60
DB      1048 CACAACACCGTGAACCTGAAGTGAACAAAGGCGGCGCCCTTCGCTCGGAGACATC 989
QY      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisArgAlaAspIlePro 80
DB      988 CTGTCCCCCAAGTTCAGTACCGGCTCAAGGTGTCGTAAGACACCCGCGACATCCCC 929
QY      81 AspTyrLyLyLeuSerPheProGluGlyPheLySerTyrGluArgValMetAsnPheGlu 100
DB      928 GACTACAAAGAGCTGTCTTCCCGAGGCGCTTCAAGTGGAGCGCGGTATGAATCTTCAG 869
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      868 GACGGCGCGCGTGGCAACCGTGAACCAAGATCTCTCCGCGAGACGGCTGCTTCACTAC 809
QY      121 LyValLyPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySerThr 140
DB      808 AAGGTGAAGTTCATCGCGGTGAATCTTCCCTCCGACGGCCCGGTATGCAAGAGAAC 749
QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProAlaGlyArgValLeuLyGlyGlu 160
DB      748 ATGGCTGGAGGCGCTTCAACCGAGGCGCTGTACCCCGACGAGGCGTCTTGAAGGCGAG 689
QY      161 IleHisLyValLeuLyLeuLyAspGlyHisTyrLeuValGluPheLySerIle 180
```

## ORIGIN

```
Alignment Scores:
Pred. No.:      1.93e-110      Length:      7616
Score:          1186.00      Matches:      219
Percent Similarity: 98.22%      Conservative: 2
Best Local Similarity: 97.33%      Mismatches: 4
Query Match:    97.69%      Indels:      0
DB:             6      Gaps:      0
```

US-10-006-922a-12 (1-225) x CQ849511 (1-7616)

```
QY      1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20
DB      3606 ATGGCTCTCTCCGGAACAGTCATCAACGAGTTCATGCTTCAAGGTGCGCATGAGAGGC 3665
QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB      3666 ACCGTGAACGGCCAGATTCAGATCGAGGCGAGGCGCGCCCTTCAAGAGGCGC 3725
QY      41 HisAsnThrValLyLeuLyValThrLyGlyGlyProLeuProPheAlaTyrAspIle 60
DB      3726 CACAACACCGTGAACCTGAAGTGAACCAAGGCGGCGCCCTTCGCTCGGAGACATC 3785
QY      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisArgAlaAspIlePro 80
DB      3786 CTGTCCCCCAAGTTCAGTACCGGCTCAAGGTGTCAGTGAAGACACCCGCGACATCCCC 3845
QY      81 AspTyrLyLyLeuSerPheProGluGlyPheLySerTyrGluArgValMetAsnPheGlu 100
DB      3846 GACTACAAAGAGCTGTCTTCCCGAGGCGCTTCAAGTGGAGCGCGGTATGAATCTTCAG 3905
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      3906 GACGCGCGCGTGGCAACCGTGAACCAAGATCTCTCCCTGCGAGGCGGCTGCTTCACTAC 3965
```

QY 121 LysValIysPheIleGlyValAsnProSerAspGlyProValMetGlnLysIleThr 140  
DB 3966 AAGGGAAAGTTCATCGGCGTGAACCTCCCTCCGACGGCCCGGTGATGACAAAGAGACC 4025  
QY 141 MetGlyTrpGluIAspSerThrGluArgLeuTyProArgAspGlyValLeuLysGlu 160  
DB 4026 ATGGCTCGGAGGCTCCACCGAGGCTGTACCCCGACGGCGCTGTAAAGGCGAG 4085  
QY 161 IleHisValAlaLeuLysLeuLysAspGlyGlyHisTyLeuValGluPheLysSerIle 180  
DB 4086 ACCCAAGGCGCTTGAAGCTGAAGGACCGGCGCACTACCTGTGAGATTCAAGTCCATC 4145  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyTyTyTyValAspSerLysLeuAsp 200  
DB 4146 TACATGGCCAAAGACCCGTGACGTGCCCGGCTACTACTACGTGAGACGCCAAGCTGAC 4205  
QY 201 IleThrSerHisAsnGluAspTyThrIleValGluGlnTyArgGluArgThrGluGlyArg 220  
DB 4206 ATCACTCTCCCAACAGAGATACACATCGTGAGAGATAGAGCGACCGAGGGCGGC 4265  
QY 221 HisHisLeuPheLeu 225  
DB 4266 CACCACTGTCTCTG 4280

RESULT 43  
LOCUS CS018252/c 9468 bp DNA linear PAT 23-FEB-2005  
DEFINITION Sequence 19 from Patent WO2005012534.  
ACCESSION CS018252  
VERSION CS018252.1 GI:60220033  
KEYWORDS

SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Alpay, L.  
TITLE Expression systems for insect pest control  
JOURNAL Patent: WO 2005012534-A 19 10-FEB-2005;  
Oxitec Limited (GB)

FEATURES  
source location/Qualifiers  
1..9468  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA710"

ORIGIN

Alignment Scores:  
Pred. No.: 2,526-110 Length: 9468  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CS018252 (1-9468)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 8727 ATGGCTCTCTCCGAGAACTCATACCGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 8668  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyValArgProTyArgGluGly 40  
DB 8667 ACCGTAAACGGCCACAGATTGAGATCGAGGGCGAGGGCGAGGGCCGCTTACGAGGGC 8608  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 8607 CACAACACCGTGAAGCTGAGAGTGAACCAAGGCGGCGCCCTGCTTCCGCTGAGATC 8548  
QY 61 LeuSerProGlnPheGlnTyArgLysValTyValLysHisProAlaAspIlePro 80  
DB 8547 CTGTCCCGCCAGTTCCAGTACGGCTCCAGAGTGAAGCAACCCCGCGACATCCCC 8488

QY 81 AspTyTrpLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 8487 GACTTACAAAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCCGCTGATGAATTGAG 8428  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTy 120  
DB 8427 GACGCGCGCGTGGAGCGATGATACCAAGATCTCTCCGTGAGAGCGGCTGCTCATCTAC 8368  
QY 121 LysValIysPheIleGlyValAsnProSerAspGlyProValMetGlnLysIleThr 140  
DB 8367 AAGGTGAAGTTCAATCGGCGTGAACCTCCCTCCACCGCCCGGTGAATGACAAAGAGACC 8308  
QY 141 MetGlyTrpGluIAspSerThrGluArgLeuTyProArgAspGlyValLeuLysGlu 160  
DB 8307 ATGGCTCGGAGGCTTCCACCGAGCGCTGTACCCCGACGGCGCTGTAAAGGCGAG 8248  
QY 161 IleHisValAlaLeuLysLeuLysAspGlyGlyHisTyLeuValGluPheLysSerIle 180  
DB 8247 ACCCAAGGCGCTTGAAGCTGAAGGACGGCGCACTACCTGTGAGATTCAAGTCCATC 8188  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyTyTyTyValAspSerLysLeuAsp 200  
DB 8187 TACATGGCCAAAGACCCGTGACGTGCCCGGCTACTACTACGTGAGACGCCAAGCTGAC 8128  
QY 201 IleThrSerHisAsnGluAspTyThrIleValGluGlnTyArgGluArgThrGluGlyArg 220  
DB 8127 ATCACTCTCCCAACAGAGATACACATCGTGAGAGATAGAGCGACCGAGGGCGGC 8068  
QY 221 HisHisLeuPheLeu 225  
DB 8067 CACCACTGTCTCTG 8053

RESULT 44  
LOCUS CS018253/c 10140 bp DNA linear PAT 23-FEB-2005  
DEFINITION Sequence 20 from Patent WO2005012534.  
ACCESSION CS018253  
VERSION CS018253.1 GI:60220034  
KEYWORDS

SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Alpay, L.  
TITLE Expression systems for insect pest control  
JOURNAL Patent: WO 2005012534-A 20 10-FEB-2005;  
Oxitec Limited (GB)

FEATURES  
source location/Qualifiers  
1..10140  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA928"

ORIGIN

Alignment Scores:  
Pred. No.: 2,746-110 Length: 10140  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CS018253 (1-10140)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 8727 ATGGCTCTCTCCGAGAACTCATACCGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 8668  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyValArgProTyArgGluGly 40  
DB 8667 ACCGTAAACGGCCACAGATTGAGATCGAGGGCGAGGGCGAGGGCCGCTTACGAGGGC 8608

QY 41 H1sAenThrVal1yLeuLeuVal1Thr1ySg1yG1yProLeuProPheAlaTrrAsp1le 60  
| | | | |  
DB 8667 CACAAACCGTGAAGCTGAAGTGAACCAAGGCGGCCCCCTTGGCTTGGGAACTC 8548  
QY 61 LeuSerProGlnPheGlnTyrg1ySer1ySVal1Tyrg1yVal1ySh1sProAlaAsp1lePro 80  
| | | | |  
DB 8547 CTGTCCCCCAAGTTCAGATGAGGCTCCAAAGGTGATGATGAAGACCCCGCGCACTATCCC 8488  
QY 81 AepTyrg1yLeuLeuSerPheProGlu1yPhe1ySTrg1yVal1yMetAsnPheGlu 100  
| | | | |  
DB 8487 GACTACAAAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGGCTGATGAACTTCGAG 8428  
QY 101 AepG1yG1yVal1yVal1Thrg1yThrg1yAsnSerSerLeuGlnAspG1yCySPhe1leTyrg 120  
| | | | |  
DB 8427 GACGCGCGGTGGCAACCGTGAACCAAGACTCTCTCTGCAAGACGCGCTGCTTCACTAC 8368  
QY 121 LysVal1yPhe1leG1yVal1yAsnPheProSerAspG1yProVal1yMetGln1yLeu1yThr 140  
| | | | |  
DB 8367 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGCGCCCGTGTATGCAAGAAAGACC 8308  
QY 141 MetG1yTrrG1yAlaSerThrg1yVal1yLeuTyrg1yProAlaAspG1yVal1yLeu1yG1yGlu 160  
| | | | |  
DB 8307 ATGGCTGGAGGCTTCAACCAAGGCTGACCGGCTGTACCCCGGACGCGCTGTGAAGGCGAG 8248  
QY 161 ILeH1sVal1yAlaLeu1yLeu1yAsnAspG1yG1yH1sTyrg1yLeuVal1yG1yPhe1ySer1le 180  
| | | | |  
DB 8247 ACCCAAGGCGCTTGAAGCTGAAGACCGGCGGCACTACTGTGTGAAGTTCAAGTCCATC 8188  
QY 181 Tyrg1yVal1yVal1yAsnProVal1yGlnLeuProG1yTyrg1yTyrg1yVal1yAsnSer1yLeu1yAsp 200  
| | | | |  
DB 8187 TACATGGCCAAAGAACCCGTGCACTGCCGCTACTACTAGTGAAGCCCAAGGCTGAC 8128  
QY 201 ILeThrSerH1sAenG1yAspTyrg1yH1sVal1yGluGlnTyrg1yVal1yThrg1yGlu1yTyrg 220  
| | | | |  
DB 8127 ATCACTCTCCCAAGAGACTACATCATGTGAGCAGTACAGCGACCGAGGCGCGC 8068  
QY 221 H1sH1sLeuPheLeu 225  
| | | | |  
DB 8067 CACCACTGTCTCTG 8053

RESULT 45  
LOCUS CS018254 10522 bp DNA linear PAT 23-FEB-2005  
DEFINITION Sequence 21 from Patent WO2005012534.  
ACCESSION CS018254  
VERSION CS018254.1 GI:60220035  
KEYWORDS  
SOURCE .  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Alphey, L.  
TITLE Expression systems for insect pest control  
JOURNAL Patent: WO 2005012534-A 21 10-FEB-2005;  
Oxitec Limited (GB)  
FEATURES  
source location/Qualifiers  
1..10522  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA1124"

ORIGIN  
Alignment Scores:  
Pred. No.: 2.87e-110 Length: 10522  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.63% Indels: 0  
DB: Gaps: 0

US-10-006-922A-12 (1-225) x CS018254 (1-10522)

QY 1 MetArgSerSer1ySAsnVal1le1yGlu1yPheMetArgPhe1ySVal1yArgMetG1yGlu 20  
| | | | |  
DB 5028 ATGGCTCTCTCCGAAGACGTATACCGAAGTTCAATGCGCTTCAAGTGGCATGAGGGC 5087  
QY 21 ThrVal1yAsnG1yH1sGluPheGlu1leGlu1yGlu1yGlu1yArgProTyrg1yGlu 40  
| | | | |  
DB 5088 ACCGTGAACGGCCACGAGTTGTGAATGAGGCGAGGGCGAGGCGCGCTTACGAGGGC 5147  
QY 41 H1sAenThrVal1yLeuLeuVal1Thr1ySg1yG1yProLeuProPheAlaTrrAsp1le 60  
| | | | |  
DB 5148 CACAAACCGTGAAGCTGAAGTGAACCAAGGCGGCCCCCTTGGCTTGGGAACTC 5207  
QY 61 LeuSerProGlnPheGlnTyrg1ySer1ySVal1Tyrg1yVal1ySh1sProAlaAsp1lePro 80  
| | | | |  
DB 5208 CTGTCCCCCAAGTTCAGATGAGGCTCCAAAGGTGATGATGAAGACCCCGCGCACTATCCC 5267  
QY 81 AepTyrg1yLeuLeuSerPheProGlu1yPhe1ySTrg1yVal1yMetAsnPheGlu 100  
| | | | |  
DB 5268 GACTACAAAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTGTATGAACTTCGAG 5327  
QY 101 AepG1yG1yVal1yVal1Thrg1yThrg1yAsnSerSerLeuGlnAspG1yCySPhe1leTyrg 120  
| | | | |  
DB 5328 GACGCGCGGTGGCAACCGTGAACCAAGACTCTCTCTGCAAGACGCGCTTCACTTAC 5387  
QY 121 LysVal1yPhe1leG1yVal1yAsnPheProSerAspG1yProVal1yMetGln1yLeu1yThr 140  
| | | | |  
DB 5388 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGCGCCCGTGTATGCAAGAAAGACC 5447  
QY 141 MetG1yTrrG1yAlaSerThrg1yVal1yLeuTyrg1yProAlaAspG1yVal1yLeu1yG1yGlu 160  
| | | | |  
DB 5448 ATGGCTGGAGGCTTCAACCAAGGCTGACCGGCTGTACCCCGGACGCGCTGTGAAGGCGAG 5507  
QY 161 ILeH1sVal1yAlaLeu1yLeu1yAsnAspG1yG1yH1sTyrg1yLeuVal1yG1yPhe1ySer1le 180  
| | | | |  
DB 5508 ACCCAAGGCGCTTGAAGCTGAAGACCGGCGGCACTACTGTGTGAAGTTCAAGTCCATC 5567  
QY 181 Tyrg1yVal1yVal1yAsnProVal1yGlnLeuProG1yTyrg1yTyrg1yVal1yAsnSer1yLeu1yAsp 200  
| | | | |  
DB 5568 TACATGGCCAAAGAACCCGTGCACTGCCGCTACTACTAGTGAAGCCCAAGGCTGAC 5627  
QY 201 ILeThrSerH1sAenG1yAspTyrg1yH1sVal1yGluGlnTyrg1yVal1yThrg1yGlu1yTyrg 220  
| | | | |  
DB 5628 ATCACTCTCCCAAGAGACTACATCATGTGAGCAGTACAGCGACCGACCGAGGCGCGC 5687  
QY 221 H1sH1sLeuPheLeu 225  
| | | | |  
DB 5688 CACCACTGTCTCTG 5702

RESULT 46  
LOCUS CS018256/c 10786 bp DNA linear PAT 23-FEB-2005  
DEFINITION Sequence 23 from Patent WO2005012534.  
ACCESSION CS018256  
VERSION CS018256.1 GI:60220037  
KEYWORDS  
SOURCE .  
ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Alphey, L.  
TITLE Expression systems for insect pest control  
JOURNAL Patent: WO 2005012534-A 23 10-FEB-2005;  
Oxitec Limited (GB)  
FEATURES  
source location/Qualifiers  
1..10786  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA670"

ORIGIN  
Alignment Scores:  
Pred. No.: 2.95e-110 Length: 10786

Score: 1186.00 Matches: 219  
 Percent Similarity: 98.22% Conservatave: 2  
 Best Local Similarity: 97.33% Mismatches: 4  
 Query Match: 97.69% Indels: 0  
 DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x CS018256 (1-10786)

```

QY      1 MetArGSeSerLyAsnValIleLySGluphMeArGPhelyeValArgMetGluGly 20
DB      8727 ATGGCTCTCCGAGAAAGCTATCAACGAGTTCAATCGCTTCAAGTGGCGATGGAGGCG 8668
QY      21 ThrValaAngIyHISgluphEgIuileGluGlyGluGlyArgProTyrgIuGly 40
DB      8667 ACCGGAACGGCCACGAGTTCCAGATCGAGGCGGAGGCGGCGCCCTTACGAGGCG 8608
QY      41 HisAenThrValLyLeuLyValThrlYsgIyGlyProLeuProPhelAaTrpAapIle 60
DB      8607 CACAAACCCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTCGCTTCCCTGGAGCATC 8548
QY      61 LeuSerProGlnPhelGlnTyrgIySerLyValTyValLySHIaProAaAapIlePro 80
DB      8547 CTGTCCCCCAAGTTCCAGTACCGGCTTCAAGGTGATGAGACACCCCGGACATCCCC 8488
QY      81 AapTyLyLyLeuSerPheProGluGlyPheLySTrGluArgValMetAenPheGlu 100
DB      8487 GACTCAAGAAAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGGTGATGAATTCGAG 8428
QY      101 AapGlyGlyValValThrlYThrgInAapSerSerLeuGlnAapGlyCyAPheIleTy 120
DB      8427 GACGGGGGGGTGGCGACCGGTGACCCAGACTCTCTCTGACGAGACGGCTTCACTTAC 8368
QY      121 LyValLyPheIleGlyValAenPheProSerAapGlyProValMetGlnLySlyThr 140
DB      8367 AAGGTGAATTCATCGGCGTGAATCTCCCTCGGACGGCCCGGTATGCAAGAAAGACC 8308
QY      141 MetGlyTyrgIuAlaSerThrgIuArgLeuTyProArgAapGlyValLeuLySGlyGlu 160
DB      8307 ATGGGCTGGAGGCGCTCAACGAGCGCTGTACCCCGGACGCGCGGTGAAAGGCGGAG 8248
QY      161 IleHisLyAlaLeuLyLeuLyAapGlyGlyHISTyLeuValGluPheLySerIle 180
DB      8247 ACCCAACAAAGCCCTGAAGCTGAAGGAGCGGGCGCACTGATGAGATTCAAATGCATC 8188
QY      181 TyrMetAlaLyLyAapProValGlnLeuProGlyTyTyTyTyValAapSerLySleAap 200
DB      8187 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCTACTACTACGTGACCGCAAGCTGGAC 8128
QY      201 IleThrSerHisAangIuAapTyThrlIleValGluGlnTyrgIuArgThrgIuArg 220
DB      8127 ATCACTCCCAACAAGAGACTACACATGTGAGAGAGTACAGAGCGCACCGAGGGCGCG 8068
QY      221 HisHisLeuPheLeu 225
DB      8067 CACCACTGTCTCTG 8053

RESULT 47
CS018251/c 11251 bp DNA linear PAT 23-FEB-2005
LOCUS      Sequence 18 from Patent WO2005012534.
DEFINITION CS018251
ACCESSION  CS018251
VERSION     GI:60220032
KEYWORDS
SOURCE      synthetic construct
ORGANISM    other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Alpay, L.
  TITLE     Expression systems for insect pest control
  JOURNAL   Patent: WO 2005012534-A 18 10-FEB-2005;
  KEYWORDS  Oxitec Limited (GB)
  FEATURES   Location/Qualifiers
            1..11251
  
```

ORIGIN

/organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="pLac56"

#### Alignment Scores:

Pred. No.: 3,116-110 Length: 11251  
 Score: 1186.00 Matches: 219  
 Percent Similarity: 98.22% Conservatave: 2  
 Best Local Similarity: 97.33% Mismatches: 4  
 Query Match: 97.69% Indels: 0  
 DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x CS018251 (1-11251)

```

QY      1 MetArGSeSerLyAsnValIleLySGluphMeArGPhelyeValArgMetGluGly 20
DB      9191 ATGGCTCTCCGAGAAAGCTATCAACGAGTTCAATCGCTTCAAGTGGCGATGGAGGCG 9132
QY      21 ThrValaAngIyHISgluphEgIuileGluGlyGluGlyArgProTyrgIuGly 40
DB      9131 ACCGGAACGGCCACGAGTTCCAGATCGAGGCGGAGGCGGCGCCCTTACGAGGCG 9072
QY      41 HisAenThrValLyLeuLyValThrlYsgIyGlyProLeuProPhelAaTrpAapIle 60
DB      9071 CACAAACCCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTCGCTTCCCTGGAGCATC 9012
QY      61 LeuSerProGlnPhelGlnTyrgIySerLyValTyValLySHIaProAaAapIlePro 80
DB      9011 CTGTCCCCCAAGTTCCAGTACCGGCTTCAAGGTGATGAGACACCCCGGACATCCCC 8952
QY      81 AapTyLyLyLeuSerPheProGluGlyPheLySTrGluArgValMetAenPheGlu 100
DB      8951 GACTCAAGAAAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGGTGATGAATTCGAG 8892
QY      101 AapGlyGlyValValThrlYThrgInAapSerSerLeuGlnAapGlyCyAPheIleTy 120
DB      8891 GACGGGGGGGTGGCGACCGGTGACCCAGACTCTCTCTGACGAGACGGCTTCACTTAC 8832
QY      121 LyValLyPheIleGlyValAenPheProSerAapGlyProValMetGlnLySlyThr 140
DB      8831 AAGGTGAATTCATCGGCGTGAATCTCCCTCGGACGGCCCGGTATGCAAGAAAGACC 8772
QY      141 MetGlyTyrgIuAlaSerThrgIuArgLeuTyProArgAapGlyValLeuLySGlyGlu 160
DB      8771 ATGGGCTGGAGGCGCTCAACGAGCGCTGTACCCCGGACGCGCGGTGAAAGGCGGAG 8712
QY      161 IleHisLyAlaLeuLyLeuLyAapGlyGlyHISTyLeuValGluPheLySerIle 180
DB      8711 ACCCAACAAAGCCCTGAAGCTGAAGGAGCGGGCGCACTGATGAGATTCAAATGCATC 8652
QY      181 TyrMetAlaLyLyAapProValGlnLeuProGlyTyTyTyTyValAapSerLySleAap 200
DB      8651 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCTACTACTACGTGACCGCAAGCTGGAC 8592
QY      201 IleThrSerHisAangIuAapTyThrlIleValGluGlnTyrgIuArgThrgIuArg 220
DB      8591 ATCACTCCCAACAAGAGACTACACATGTGAGAGAGTACAGAGCGCACCGAGGGCGCG 8532
QY      221 HisHisLeuPheLeu 225
DB      8531 CACCACTGTCTCTG 8517

RESULT 48
CS018250/c 11570 bp DNA linear PAT 23-FEB-2005
LOCUS      Sequence 17 from Patent WO2005012534.
DEFINITION CS018250
ACCESSION  CS018250
VERSION     GI:60220031
KEYWORDS
SOURCE      synthetic construct
ORGANISM    synthetic construct
  
```

other sequences; artificial sequences.

REFERENCE	1
AUTHORS	Alpay, L.
TITLE	Expression systems for insect pest control
JOURNAL	Patent: WO 2005012334-A 17 10-FEB-2005;
FEATURES	Oxitec Limited (GB)
SOURCE	Location/Qualifiers
	1. 11570

### Alignment Scores:

Pred. No.:	3.22e-110	Length:	11570
Score:	1186.00	Matches:	219
Percent Similarity:	98.42%	Conservative:	2
Best Local Similarity:	97.33%	Mismatches:	4
Query Match:	97.69%	Indels:	0
GB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x CS018250 (1-11570)

Qy 1 MetArgSerSerLysAsnValIleIleGluPheMetArgPheLysValArgMetGluGly 20  
||| : : |||  
Db 8886 ATGCGCTCCTTCGAGAACGTATCACCAGATTATGCCGTTCAAGGTGCGCATGAGGGGC 88

QY 21 ThrValAsnGlyHtIeGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 8826 ACCGTGAACGCCACGAGTTTCGAGATCGAGCGAGGCGAGGCGCCCTTACGAGGC 87

**Oy** 41 HisaentThrValIysleuLysValThrlysgGlyProleuPProPhealAtTpaspile 60  
|||  
**Pb** 8766 CACACACCCCGAAGCTGAAGTGACCAGAAGGGGCGGCCCTGTGCCTTCGGCTGGGATC 8767

Qy 61 LeuSerProGlnPheGlnTyrGlySerIySValTyrValIySHiSProAlaAspIlePro 80  
 8706 CTGTCCTCCCTCCAGTTCAGTACGAGTCAAGAGTCACTGAGACCCGCGGACATCCCC 86

Qy 81 AspyrYrlyslvsleuserPheProGluGlyPheIystrPGIwargValmetAsnPhenGlu 10  
|||  
R646 GACCTCAAGGAAGCTGTCTTCCCCGAGGACTTCAAGTGGGAGCGGTGATGAACCTTCGAG A5  
|||

101 AspglyglyvalvalthrvalthrGlnaspserSerLeuGlnaspGlyCyspheIleTyr 12

121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr 14

QY 141 MetGlyTyrGlnAlaSerThrGlnArgLeuTyrProArgAspGlyValLeuLysGlyGlu 16

161 ILeH1sLyva1aLeu1ySLeu1ySAsp1yG1yH1sTYrLeuVa1G1uPhelySer1le 18

QY 181 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
182 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
183 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
184 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
185 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
186 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
187 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
188 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
189 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
190 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
191 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
192 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
193 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
194 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
195 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
196 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
197 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
198 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
199 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||  
200 TyrMetAlaIaYbLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysIleuAsp 20  
|||||

Qy 201 IleThrSerHisAsnGluAspPyrThrIleValGluGlnTyrGluArgThrGluGlyArg 22

```

Qy      221 HisHisLeuPheLeu 225
          |||||
Db      8226 CACCACCTGTTCCCTG 8212

```

## RESULT 49

CS018255	CS018255	11867 bp	DNA	linear	PAT 23-FEB-2005
LOCUS	Sequence 22 from Patent WO2005012534.				
DEFINITION					
ACCESSION	CS018255				
VERSION	CS018255.1	GI:60220036			

**ORGANISM** synthetic construct  
other sequences; artificial sequences.

REFERENCE  
1  
Alphay, L.  
Expression systems for insect pest control  
Patent: WO 2005/012534-A 22 10-FEB-2005;  
Oxitec Limited (GB)

source

```

/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pLA1188"

```

Alignment Scores:  
Pred. No.: 3.32e-110 Length

Score:	188.00	Matches:	219
Percent Similarity:	98.22%	Conservative:	2
Best Local Similarity:	97.33%	Mismatch:	0
Query Match:	97.69%	Indels:	0

DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x CS018255 (1-11867)			

Db 6373 ATGGCCTCTCCGAGAACGTCATCACCGAGTTCATGGCGCTTCAAGGTGCCCATGAGGGC 643

QY 21 ThxValAsnGlyIh1SGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 6433 ACCGTGAACGGCCACGAGTTGGAATCGAGGGCGAGGGCGAGGGCCGCCCTACAGAGGC 649

Qy	41	HissnThVallylsleuLysValThrlYsgLysProleupProhealatripspIle	60
Db	6493	CACACACCGTGAAGCTGAAGTGCACCAAGGGCGCCCCCTTCCTTCGCTGGGACATC	655

Qy 61 LeuSerProGlnpneGlnIArgLysSerLysValTyrValLysHisProLysAspIlePro 80  
Db 6553 CTGTCCCCCAGTTCAGTACGGGTCCAGGTGTACGTGAAGCAGCCCCGGCGACATCCCC 661

QY 81 AspPyrLrVLSLysLeuSerPhePrGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 6613 GACACAGAGAGCTCTTCCCGGAGGCTTCAAGTGGAGCGGCTGATGACACTTCGAG 6677

QY 101 AspGclgYValValThrValThnGlnAppSerSerIengGlnAspGlyCysPhenIleYr 120  
Db 6673 GACGGCGGCGTGGGACCCGTGACCCAGACTCTCTCTCTGACGAGCGCTCTTATCTAC 6733

QY 121 LytVallvPheHieGLyValaAnPhePSeSerApGLyProValMetGlnLysLysThr 140  
Db 6733 AAGGTGAAGTTCAATGGCGGTGAACCTCCCTCCGACGGCCCGTATGCGAGAGAGACC 6799

QY 141 MetGLYTPGlnLase-rThrGluArgLeuTYrProArgAspGluValLeuLysGlu 160  
Db 6793 ATGGAGCTGGAGGCTCCACCGAGCGCCTGACCCCGCAGCGGCTGCTGAGAGGGCGAG 685

QY 161 ILHLSYSAIALEuLYbEuLYbAspGlycylHISLYLeuValGluPhelysSerIle 180

Db 6853 ACCCAACAAGCCCTGAGCTGAAGGACGGCGCACTACCTGTGTGAGATTCAAGTTCATC 6911

201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGluArg 220



DB 6973 ATCACTCCCAACAGAGACTACATCGTGAGCAGTACGACGACCGAGGCGGC 7032  
QY 221 HSHSLepheleu 225  
DB 7033 CACCACTGTTCCTG 7047

RESULT 50  
CS018249 11920 bp DNA linear PAT 23-FEB-2005  
DEFINITION Sequence 16 from Patent WO2005012534.  
ACCESSION CS018249  
VERSION CS018249.1 GI:60220030  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
Alphey, L.  
AUTHORS Expression systems for insect pest control  
TITLE Patent: WO 2005012534-A 16 10-FEB-2005;  
JOURNAL Oxitec Limited (GB)  
FEATURES  
source location/Qualifiers  
1..11920  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA513"

ALIGNMENT SCORES:  
Pred. No.: 3,34e-110 Length: 11920  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CS018249 (1-11920)

QY 1 MetAAGSerSerLybAnValIleYsGluPhemeArGPhelysValaArgMetGluY 20  
DB 7953 ATGGCCTCTCCCGAAGACGTCATCACCGAGTTCATCGCTTCMAAGTGCGCATGAGGCG 8012  
QY 21 ThrValAsnGlyHfAgIupheGluIleGluYgIyGluYArpProYrYrGluY 40  
DB 8013 ACCGTAAACGCGACAGATTCCAGATCGAGGCGAGGCGGCGCCCTTCAGAGGC 8072  
QY 41 HfAsnThrValYbLeuYsValIthrYsGlyYgIyProLeuProPhelaATrPaPle 60  
DB 8073 CACAAACCGTGAAGCTGAAGTGAACCAAGGCGGCGCCCTTCGCTGAGCATC 8132  
QY 61 LeuSerProGlnPheGlnIYrGlySerLybValYrValYbHfProIaApPlePro 80  
DB 8133 CTGTCCCCCAAGTTCAGATCGGCTCCAGGTGTACGTGAAGACCCCGCGCATCCCC 8192  
QY 81 AspTyrLybYbLeuSerPheProGluYgIyPheLybTrpGluYrValaMetAanPheGlu 100  
DB 8193 GACTTAAAGAGGCTGTCTTCCCGAGGGCTTCAGATGGAGGCGGTGATGAATTCGAG 8252  
QY 101 AspGlyGlyValaValaThrValIthrGlnAaspSerSerLeuGlnaApGlyCybPheIleYr 120  
DB 8253 GACGGGGCGGTGGAGACCGTGAACCGAGACTCCCTCCGAGGCGGCTGCTCATCTAC 8312  
QY 121 LysValaLybPheIleGlyValaAanPheProSerPheGlyProValaMetGlnYbYbThr 140  
DB 8313 AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGAGCGGCGCGGTGAGAGAAAGAAC 8372  
QY 141 MetGlyTTrpGluYbAsnThrGluYrGluYrProArgApGlyValaLeuYsGlyGlu 160  
DB 8373 ATGGAGCGGAGGCTCAACGAGCGGCTGTACCCCGCGAGCGGCTGTGAGAGGAGAG 8432  
QY 161 HfHfIleYbAlaLeuYbLeuYbAaspGlyHfIleYrLeuValaGluPheLybSerIle 180

DB 8433 ACCCAAGAGCCCTGAACTAAAGACGCGCCCATCACTCGTGGAGTTCAAGTCCATC 8492  
QY 181 TyrMetAlaYbYsProValaGluProGlyYrYrYrYrValaAaspSerLybLeuAap 200  
DB 8493 TACATGGCCMAAGAGCCGCTGACGCTGCCGCTTACTACTAGTGAACGCCAAGCTGAGC 8552  
QY 201 HfThrSerHfAsnGluYbApTyrThrIleValaGluGlnYrGluYrGluYrGluYrArg 220  
DB 8553 ATCACTCCCAACAGAGACTACATCGTGAGCAGTACGAGCGACCGAGGCGCGC 8612  
QY 221 HSHSLepheleu 225  
DB 8613 CACCACTGTTCCTG 8627

RESULT 51  
CS018257 14720 bp DNA linear PAT 23-FEB-2005  
DEFINITION Sequence 24 from Patent WO2005012534.  
ACCESSION CS018257  
VERSION CS018257.1 GI:60220038  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.

REFERENCE 1  
Alphey, L.  
AUTHORS Expression systems for insect pest control  
TITLE Patent: WO 2005012534-A 24 10-FEB-2005;  
JOURNAL Oxitec Limited (GB)  
FEATURES  
source location/Qualifiers  
1..14720  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA1038"

ALIGNMENT SCORES:  
Pred. No.: 4,32e-110 Length: 14720  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CS018257 (1-14720)

QY 1 MetAAGSerSerLybAnValIleYsGluPhemeArGPhelysValaArgMetGluY 20  
DB 7382 ATGGCCTCTCCCGAAGACGTCATCACCGAGTTCATCGCTTCMAAGTGCGCATGAGGCG 7323  
QY 21 ThrValAsnGlyHfAgIupheGluIleGluYgIyGluYArpProYrYrGluY 40  
DB 7322 ACCGTAAACGCGACAGATTCCAGATCGAGGCGAGGCGGCGCCCTTCAGAGGC 7263  
QY 41 HfAsnThrValYbLeuYsValIthrYsGlyYgIyProLeuProPhelaATrPaPle 60  
DB 7262 CACAAACCGTGAAGCTGAAGTGAACCAAGGCGGCGCCCTTCGCTGAGCATC 7203  
QY 61 LeuSerProGlnPheGlnIYrGlySerLybValYrValYbHfProIaApPlePro 80  
DB 7202 CTGTCCCCCAAGTTCAGATCGGCTCCAGGTGTACGTGAAGACCCCGCGCATCCCC 7143  
QY 81 AspTyrLybYbLeuSerPheProGluYgIyPheLybTrpGluYrValaMetAanPheGlu 100  
DB 7142 GACTTAAAGAGGCTGTCTTCCCGAGGGCTTCAGATGGAGGCGGTGATGAATTCGAG 7083  
QY 101 AspGlyGlyValaValaThrValIthrGlnAaspSerSerLeuGlnaApGlyCybPheIleYr 120  
DB 7082 GACGGGGCGGTGGAGACCGTGAACCGAGACTCCCTCCGAGGCGGCTGTGAGAGGAG 7023  
QY 121 LysValaLybPheIleGlyValaAanPheProSerPheGlyProValaMetGlnYbYbThr 140

```

Db      7022 AAGGGAAGTTCATGGGGTGAACCTCCCTCCGAGCGGCCCGCTGATGACAGAAAGACC 6963
Oy      141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyCln 160
Db      6962 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGGGCTGTCAAGGGCGAG 6903
Oy      161 ILeHslyAlaLeuIysLeuIysAspGlyGlyHsTyrLeuValGluPheIysSerIle 180
Db      6902 ACCCAAGAGCCCTTAAGCTGAAGGACGAGCGGCACTACCTGGTGGAGTTCAAGTTCATC 6843
Oy      181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200
Db      6842 TACATGGCCAGAAAGCCCGTGCAGCTGCCCGCTACTACTAGTGAACGCCAAGCTGGAC 6783
Oy      201 ILeHsSerHsAsnGluAspTyrThrIleValGlnGlnTyrGluArgThGlnGlyArg 220
Db      6782 ATCACTCCCAACAAGAGACTACCATCGTGAAGCAGTACGAGCGCACCGAGGGCCGC 6723
Oy      221 HsHsIleuPheLeu 225
Db      6722 CACCACTGTTCTTG 6708

RESULT 52
LOCUS      CQ981073              16157 bp    DNA          linear    PAT 25-JAN-2005
DEFINITION Sequence 3 from Patent WO2005003364.
ACCESSION  CQ981073
VERSION     CQ981073.1  GI:58190337
KEYWORDS    .
SOURCE      synthetic construct
ORGANISM    synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Alphey, L.
  TITLE     Stable integrands
  JOURNAL   Patent: WO 2005003364-A 3 13-JAN-2005;
            Oxitec Limited (GB)
FEATURES    Location/Qualifiers
            source          1..16157
                           /organism="synthetic construct"
                           /mol_type="unassigned DNA"
                           /db_xref="taxon:32630"
                           /note="predicted sequence of pUA1025 construct"
ORIGIN
Alignment Scores:
Pred. No.:      4,83e-110      Length:      16157
Score:          1186.00        Matches:      219
Percent Similarity: 98.22%      Conservative: 2
Best Local Similarity: 97.33%    Mismatches:  4
Query Match:    97.69%         Indels:      0
DB:             Gaps:         0

US-10-006-922a-12 (1-225) x CQ981073 (1-16157)
Oy      1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGlnGly 20
Db      14524 ATGGCTCTCTCCGAAACGTCATCCACGATTCACGCTTCACAGTGCAGATGGAGGGC 14465
Oy      21 ThrValAsnGlyHsGluPheGluIleGluGlyGlnGlyArgProTyrGlyGly 40
Db      14464 ACCGTGAACGGCCACGATTCGAGATCGAGGCGAGGCGGAGGCCGCCCTTACGAGGGC 14405
Oy      41 HsAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTTPAspIle 60
Db      14404 CACAACACCGTAGAGCTGAAGGTGACCAAGGGCGGCCCTGCGCTTGCGGTGGACATC 14345
Oy      61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHsIleProAlaAspIlePro 80
Db      14344 CTGTCCCTCCCAAGTTCAGATCAGGCTCCCAAGGTGATCGTGAAGCAACCCGCGACATCCCC 14285
Oy      81 AspTyrIysIysLeuSerPheProGluGlyPheIysTyrGluArgValMetAsnPheGlu 100

```

```

Db      14284 GACTACAAAGAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGATGATGAATCTGAG 14225
Oy      101 AspGlyGlyValIleThrValThrGlnAspSerSerIleGlnAspGlyCysPheIleTyr 120
Db      14224 GACCGCGCGCGGGACCGTGAACCAAGACTCTCCCTCGAGGACGGCTGCTTACTATC 14165
Oy      121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysLeuThr 140
Db      14164 AAGGTGAAGTTCACTGGCGTGAATCTTCCCTCCAGCGGCCCGCTGATGCAAGAAAGACC 14105
Oy      141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyCln 160
Db      14104 ATGGGCTGGAGGCTCTCACAGAGCGCTGTACCCCGCGAGCGGCGTGTGAAGGGCGAG 14045
Oy      161 ILeHslyAlaLeuIysLeuIysAspGlyGlyHsTyrLeuValGluPheIysSerIle 180
Db      14044 ACCCAAGAGCCCTGAAGCTGAAGAGAGCGGGCCACTTCTGTGAGTTCAAGTTCATC 13985
Oy      181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200
Db      13984 TACATGGCCAGAAAGCCCGTGCAGCTGCCCGCTACTACTAGTGAACGCCAAGCTGGAC 13925
Oy      201 ILeHsSerHsAsnGluAspTyrThrIleValGlnGlnTyrGluArgThGlnGlyArg 220
Db      13924 ATCACTCCCAACAAGAGACTACCATCGTGAAGCAGTACGAGCGCACCGAGGGCCGC 13865
Oy      221 HsHsIleuPheLeu 225
Db      13864 CACCACTGTTCTTG 13850

RESULT 53
LOCUS      CQ981074              16157 bp    DNA          linear    PAT 25-JAN-2005
DEFINITION Sequence 4 from Patent WO2005003364.
ACCESSION  CQ981074
VERSION     CQ981074.1  GI:58190338
KEYWORDS    .
SOURCE      synthetic construct
ORGANISM    synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Alphey, L.
  TITLE     Stable integrands
  JOURNAL   Patent: WO 2005003364-A 4 13-JAN-2005;
            Oxitec Limited (GB)
FEATURES    Location/Qualifiers
            source          1..16157
                           /organism="synthetic construct"
                           /mol_type="unassigned DNA"
                           /db_xref="taxon:32630"
                           /note="predicted sequence of pUA1125 construct"
ORIGIN
Alignment Scores:
Pred. No.:      4,83e-110      Length:      16157
Score:          1186.00        Matches:      219
Percent Similarity: 98.22%      Conservative: 2
Best Local Similarity: 97.33%    Mismatches:  4
Query Match:    97.69%         Indels:      0
DB:             Gaps:         0

US-10-006-922a-12 (1-225) x CQ981074 (1-16157)
Oy      1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGlnGly 20
Db      14524 ATGGCTCTCTCCGAAACGTCATCCACGATTCATGCGCTTCACAGTGCAGATGGAGGGC 14465
Oy      21 ThrValAsnGlyHsGluPheGluIleGluGlyGlnGlyArgProTyrGlyGly 40
Db      14464 ACCGTGAACGGCCACGATTCGAGATCGAGGCGAGGCGGAGGCCGCCCTTACGAGGGC 14405
Oy      41 HsAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTTPAspIle 60

```

```

Db      14404  CACAAACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCTCCCTGCGCGGAGCATC 14345
QY      61      LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysPheProAlaAspIlePro 80
Db      14344  CTGTCCTCCCACTTCAGTACGCGCTCCAGAGGTGACGTGACACCCCGCCACATCTCCC 14285
QY      81      AspTyrIysLysLeuSerPheProGlnGlyPheIleTyrGlnArgValMetAsnPheGlu 100
Db      14284  GACTACAAAGAGCTGCTCTCCCGAGGCTTCAGGTGGAGCGGTGATGAACTTCGAG 14225
QY      101     AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      14224  GACGCGGCGGTGCGACCGGTGACCCAGGACTCTCTCCGACAGAGCGGCTGCTTCACTAC 14165
QY      121     LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr 140
Db      14164  AAGGTGAAGTTCATCGCGCTGAACTTCCCTCCGACGCGCCCGTGAAGCAAGAACCC 14105
QY      141     MetGlyTyrPGLuAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db      14104  ATGGCTGGGAGGCTCCACCGAGCCCTGTACCCCGGAGCGGCTGTGAAAGGCGAG 14045
QY      161     IleHisLysValLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db      14044  ACCCAAGAGCCCTGAAGCTGAAGGACGCGGCGCACTACCTGGTGAAGTTCAAGTTCATC 13985
QY      181     TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db      13984  TACATGGCAAGAGCCGCTGACGTGCGCGGCTACTACTGAGACCGCAAGCTTGAGC 13925
QY      201     IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGlnLysArgThrGlnGlyArg 220
Db      13924  ATCACTCTCCCAACAAGAGACTACCACTGTGAGACAGTACAGGACGCCAGGCGGCCCG 13865
QY      221     HisHisLeuPheLeu 225
Db      13864  CACCACTGTCTCTG 13850

RESULT 54
AX686888  LOCUS      AX686888      898 bp      DNA      linear      PAT 31-MAR-2003
DEFINITION Sequence 11 from Patent WO0127150.
ACCESSION AX686888
VERSION    AX686888.1 GI:29409468
KEYWORDS
SOURCE     Discosoma sp.
ORGANISM   Discosoma sp.
            Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;
            Corallimorpharia; Discosomacidae; Discosoma.
REFERENCE  1      Lukyanov, S.A., Pradkov, A.F., Labas, Y.A., Matz, M.V. and Terzikikh, A.
            Anthozoa derived chromo/fluoroproteins and methods for using the
            same
JOURNAL    Patent: WO 0127150-A 11 19-APR-2001;
            Clontech Laboratories Inc. (US)
FEATURES   source
            1..898 /organism="Discosoma sp."
            /mol_type="unassigned DNA"
            /db_xref="taxon:86600"

ORIGIN
Alignment Scores:
Pred. No.:      2,28e-111      Length:      898
Score:          1184.00        Matches:      219
Percent Similarity: 98.22%      Conservative: 2
Best Local Similarity: 97.33%      Mismatches: 4
Query Match:    97.53%      Indels:      0
DB:             6              Gaps:        0
US-10-006-922a-12 (1-225) x AX686888 (1-898)

```

```

QY      1      MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGly 20
Db      93      ATGAGGTCTTCCAGAAAGTATACAGAGATTCATGAGTTTAAGTTCCGATGGAAAGA 152
QY      21      ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGlyGly 40
Db      153      ACGGTCAATGGGCAAGATTTGAATATAGAGCGAAGGAGGAGGCCATACGAAGGC 212
QY      41      HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60
Db      213      CACATATCCGTAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGCTGGAGATTT 272
QY      61      LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysPheProAlaAspIlePro 80
Db      273      TTGTACACACAAATTCAGATATGAAGCAAGGTATATGTCCAGACCTTCGCCACATACCA 332
QY      81      AspTyrIysLysLeuSerPheProGlnGlyPheIleTyrGlnArgValMetAsnPheGlu 100
Db      333      GACTATTAAGAGCTGTCTATCTCGAAGATTAAATGGAAAGGCTCATGAACTTTGAA 392
QY      101     AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      393      GACGGTGGCGTCTTACTGTAAACCGAGATTCAGTTGACAGATGGCTGTTCATCTAC 452
QY      121     LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr 140
Db      453      AAGTCAGTTCAATGCGCTGTAACCTTCTCCGATGGACCTGTATATGCAAAAGAGACA 512
QY      141     MetGlyTyrPGLuAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db      513      ATGGCTGGGAGAGCCAGACACTGAGCGTTGTATCTCGATGAGGCGTGTGAAGAGAG 572
QY      161     IleHisLysValLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db      573      ATTCATTAAGGCTCTGAAGCTGAAGACGCGTGTCTATCTTGAATTTGAATCAAAAGTATT 632
QY      181     TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db      633      TACATGGCAAGAGCGCTGACGCTACCGAGGATCTACTATGTTGACTCCAAACTGGAT 692
QY      201     IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGlnLysArgThrGlnGlyArg 220
Db      693      ATTAACAGCCACAAGACATATCAATCGTTGAGCATATGAAGAAAGAACCGAGGAGCC 752
QY      221     HisHisLeuPheLeu 225
Db      753      CACCACTGTCTCTT 767

RESULT 55
AY916793  LOCUS      AY916793      3441 bp      DNA      circular SYN 30-MAR-2005
DEFINITION Dual fluorescent protein cloning vector pGRFP, complete sequence.
ACCESSION AY916793
VERSION    AY916793.1 GI:60101687
KEYWORDS
SOURCE     Dual fluorescent protein cloning vector pGRFP
ORGANISM   Dual fluorescent protein cloning vector pGRFP
            other sequences; artificial sequences; vectors.
REFERENCE  1      (bases 1 to 3441)
            Choe, J., Guo, H.H. and van den Engn, G.
            A dual-fluorescence reporter system for high-throughput clone
            characterization and selection by cell sorting
            (ex) Nucleic Acids Res. 33 (5), E49 (2005)
JOURNAL    PUBMED
            15767274
AUTHORS    Choe, J., Guo, H.H. and van den Engn, G.
TITLE      Direct Submission
JOURNAL    Submitted (01-FEB-2005) Laboratory of Ger van den Engn, Institute
            for Systems Biology, 1441 N 34th St, Seattle, WA 98103-8904, USA
FEATURES   Location/Qualifiers
            1..3441
            /organism="Dual fluorescent protein cloning vector pGRFP"
            /mol_type="other DNA"

```







ORGANISM Red H-Pelican DsRed.T4 transformation vector  
REFERENCE other sequences; artificial sequences; vectors.  
AUTHORS 1 (bases 1 to 10141)  
TITLE Barolo,S. and Posakony,J.W.  
ADVANCES Advances in Drosophila Transgenic Reporters I: Fast-Maturing RFP  
(DsRed.T4) and Nuclear RFP in Insulated Vectors  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 10141)  
AUTHORS Barolo,S. and Posakony,J.W.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUN-2003) Division of Biology/CDB, UCSD, 4121 Bonner  
Hall, MC 0349, La Jolla, CA 92093-0349, USA  
FEATURES  
source  
1. 10141  
/organism="Red H-Pelican DsRed.T4 transformation vector"  
/mol\_type="other DNA"  
/db\_xref="taxon:240544"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,82e-108 Length: 10141  
Score: 1165.00 Matches: 216  
Percent Similarity: 97.33% Conservative: 3  
Best Local Similarity: 96.00% Mismatches: 6  
Query Match: 95.96% Indels: 0  
Gaps: 0  
US-10-006-922A-12 (1-225) x AY342347 (1-10141)  
QY 1 MetAAGSerSerLybAenVal11elysGluPhemeArzPhelybValAArgMetGluGly 20  
DB 285 ATGGCCTCTCCGAGAGACGTCATCAAGAGGTTCACTCCGTTCAAGGTGCGCATGGAGGCG 344  
QY 21 ThrValAenGlyH1egluPheliuegluGluGlyGluGlyYArgProTyrgluGly 40  
DB 345 TCCGTGAACGGCCACAGTTCCAGATCGAGGCGAGGCGGCGGCGCCCTTCAGAGGCG 404  
QY 41 H1AenThrValLybLeuLybValThrlYsglyGlyProLeuProBhealAtrPaapile 60  
DB 405 ACCCAGACCGCCAGCTGAAGGTGACCAAGGGGCGCCCTTCGCTGGAGACATC 464  
QY 61 LeuSerProGlnPheliuegluGlySerLybValTyValLybH1AenProAlaapilPero 80  
DB 465 CTGTCCCCCAGTTCCAGTACGAGCTCCAGGTGTACGTAGACACCCCGCCGACATCCGCC 524  
QY 81 AapTyLybLybLeuSerPheProgluGlyPheLybTrrgluArgValMetAenPheglu 100  
DB 525 GACTACAAAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTATGAACTTCGAG 584  
QY 101 AapGlyGlyValValThrlValThrglnAapSerSerLeuGlnAapGlyCySpheileTy 120  
DB 585 GACGGCGCGTGTGACCGTGAACCGAGACTCTCTCCGTCAGAGCGGCTTCATCTAC 644  
QY 121 LybValLybPhelelleglyValAenPheProSerAapGlyProValMetGlnLybTytr 140  
DB 645 AAGGTAAAGTTCACTCGCGTGAACCTTCCCTCCGACGCGCCCGTAAATGACAGAAAGACT 704  
QY 141 MetGlyTrrgluA1AserThrgluArgLeuTyProArgAapGlyValLeuLybGlyglu 160  
DB 705 ATGGGTGGAGGCTTCACCGAGCGCTGTACCCCGCGACGCGGCTGAAAGGGGAG 764  
QY 161 H1eH1eLybValAenLybLeuLybAapGlyGlyH1eTyLeuValGluPhelybSerile 180  
DB 765 ATCCACAAAGCCCTGAAGCTGAAGAGCGGCGGCACTACTGGTGGAGTTCAAGTCCATC 824  
QY 181 TyMetAlaLybLybProValGlnLeuProGlyTyTyTyTyValAapSerLybLeuAap 200  
DB 825 TACATGGCCAAAGAGCCCGTGAAGCTGCGCGGCTACTACTACGTGAGCTCAAGAGCTGAG 884  
QY 201 H1eThSerH1eAenGlyuAapTyTrThrlValGluGlnTyrgluArgThrgluGlyArg 220  
DB 885 ATCACTCCCAACAGAGACTACATCATGTGAGACATACAGACCGCGCGAGGCGCGC 944

QY 221 H1eH1eLeuPheleu 225  
DB 945 CACACCTGTTCCTG 959  
RESULT 60  
LOCUS AY342348  
DEFINITION Red H-Stinger DsRed.T4-NLS transformation vector, complete  
ACCESSION AY342348  
VERSION AY342348  
KEYWORDS AY342348.1 GI:33358310  
SOURCE  
ORGANISM Red H-Stinger DsRed.T4-NLS transformation vector  
REFERENCE Red H-Stinger DsRed.T4-NLS transformation vector  
AUTHORS other sequences; artificial sequences; vectors.  
TITLE 1 (bases 1 to 10276)  
JOURNAL Advances in Drosophila Transgenic Reporters I: Fast-Maturing RFP  
[DsRed.T4] and Nuclear RFP in Insulated Vectors  
REFERENCE 2 (bases 1 to 10276)  
AUTHORS Barolo,S. and Posakony,J.W.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUN-2003) Division of Biology/CDB, UCSD, 4121 Bonner  
Hall, MC 0349, La Jolla, CA 92093-0349, USA  
FEATURES  
source  
1. 10276  
/organism="Red H-Stinger DsRed.T4-NLS transformation  
vector"  
/mol\_type="other DNA"  
/db\_xref="taxon:240545"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,88e-108 Length: 10276  
Score: 1165.00 Matches: 216  
Percent Similarity: 97.33% Conservative: 3  
Best Local Similarity: 96.00% Mismatches: 6  
Query Match: 95.96% Indels: 0  
Gaps: 0  
US-10-006-922A-12 (1-225) x AY342348 (1-10276)  
QY 1 MetAAGSerSerLybAenVal11elysGluPhemeArzPhelybValAArgMetGluGly 20  
DB 288 ATGGCCTCTCCGAGAGACGTCATCAAGAGGTTCACTCCGTTCAAGGTGCGCATGGAGGCG 347  
QY 21 ThrValAenGlyH1egluPheliuegluGluGlyGluGlyYArgProTyrgluGly 40  
DB 348 TCCGTGAACGGCCACAGTTCCAGATCGAGGCGAGGCGGCGGCGCCCTTCAGAGGCG 407  
QY 41 H1AenThrValLybLeuLybValThrlYsglyGlyProLeuProBhealAtrPaapile 60  
DB 408 ACCCAGACCGCCAGCTGAAGGTGACCAAGGGGCGCCCTTCGCTGGAGACATC 467  
QY 61 LeuSerProGlnPheliuegluGlySerLybValTyValLybH1AenProAlaapilPero 80  
DB 468 CTGTCCCCCAGTTCCAGTACGAGCTCCAGGTGTACGTAGACACCCCGCCGACATCCGCC 527  
QY 81 AapTyLybLybLeuSerPheProgluGlyPheLybTrrgluArgValMetAenPheglu 100  
DB 528 GACTACAAAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTATGAACTTCGAG 587  
QY 101 AapGlyGlyValValThrlValThrglnAapSerSerLeuGlnAapGlyCySpheileTy 120  
DB 588 GACGGCGCGTGTGACCGTGAACCGAGACTCTCTCCGTCAGAGCGGCTTCATCTAC 647  
QY 121 LybValLybPhelelleglyValAenPheProSerAapGlyProValMetGlnLybTytr 140  
DB 648 AAGGTAAAGTTCACTCGCGTGAACCTTCCCTCCGACGCGCCCGTAAATGACAGAAAGACT 707  
QY 141 MetGlyTrrgluA1AserThrgluArgLeuTyProArgAapGlyValLeuLybGlyglu 160

D	b	708	ATGGAGCTGGGAGACCCCTTCACCGAGCCGCTGTACCCCGCAGCGGCTGCTGAAGGGGAG	767
Q	y	161	lethlsbyslalaleuylsleuylsaspqlvglyhlslyrleuvalgluphelseserile	180
D	b	768	ATCCACAAGGCCCTGAAAGCTGAAGAGACGGCGGCACCTACCTGTGGAGTTCAAGTTCATC	827
Q	y	181	TyrMetkAlaIaIysIyProValGlnleuProGlyIyTyrTyrTyrValaIaspSerIyLeuasp	200
D	b	828	TACATGGCCAGGAAGCCCGTGCAGCTCGCCCGCTACTACTACGTGAGACTCCAGACTGGAC	887
Q	y	201	lethserhsaenglsaeplyrthrilevalgluglnlyrglwaacthgluglyarg	220
D	b	888	ATCACCTTCCACACAGAGACTACACCATCTGTAGACAGTACGAGCGCCGAGAGGCCGC	947
Q	y	221	HlsHlsleuPhelenu 225	
D	b	948	CACCACCTGTTCTCG 962	
RESULT 61				
LOCUS		AY490568	10481 bp	DNA linear SYN 24-DEC-2003
DEFINITION		UNS-Red Stinger DsRed.T4-NLS Drosophila	UNS-RF transformation	
ACCESSION		AY490568		
VERSION		AY490568.1	GI:40218087	
KEYWORDS				
SOURCE		UNS-Red Stinger DsRed.T4-NLS Drosophila	UNS-RF transformation	
ORGANISM		vector	UNS-Red Stinger DsRed.T4-NLS Drosophila	UNS-RF transformation
REFERENCE		other sequences; artificial sequences; vectors.		
AUTHORS		1 (bases 1 to 10481)		
TITLE		Barolo,S. and Posakony,J.W.		
JOURNAL		Advances in Drosophila Transgenic Reporters I: Fast-Maturing RFP		
REFERENCE		[Dered.T4] and Nuclear RFP in Insulated Vectors		
AUTHORS		Unpublished		
TITLE		2 (bases 1 to 10481)		
JOURNAL		Barolo,S. and Posakony,J.W.		
REFERENCE		Direct Submission		
AUTHORS		Submitted (13-NOV-2003) Cell & Developmental Biology, University of		
TITLE		Michigan Medical School, 5732 Med Sci 2, Ann Arbor, MI 48109, USA		
JOURNAL		Location/Qualifiers		
FEATURES		1. .10481		
SOURCE		/organism="UNS-Red Stinger DsRed.T4-NLS Drosophila		
ORIGIN		/transformation vector"		
		/mol_type="other DNA"		
		/db_xref="taxon:258595"		
Alignment Scores:				
Pred. No.:		3.98e-108	Length:	10481
Score:		1165.00	Matches:	216
Percent Similarity:		97.33%	Conservative:	3
Best Local Similarity:		96.00%	Mismatches:	6
Query Match:		95.96%	Indels:	0
DB:		11	Gaps:	0
US-10-006-922A-12		(1-225) x AY490568	(1-10481)	
Q	y	1	MetArgSerSerIyBaHValIleIygluPhMetArgPhelysValArgMetGlugly	20
D	b	494	ATGGCCTTCCTCCGAGGACGTCATCAAGAGTTCATGCGCTTCAAGGTCGCGCATGAGAGGC	553
Q	y	21	ThrValaIaIaIyHlsGluPhelGlnleGlnIyglIuIyglIuIyglIyArgProTyrGlugly	40
D	b	554	TCCCTGAAACGGCCACGAGTTTGAAGTTCAGAGGCGCAGAGGCGCGCCCTTACGAGGC	613
Q	y	41	HlsaHnThrValIyIyLeuIyValThrIyIyGlyIyProIeuProPhaIaIaTyrAspIle	60
D	b	614	ACCCAGACCGCGCAAGCTGAAGGTACCAAGGCGCGCCCTTGCCCTTGAGGACATC	673
Q	y	61	LeuSerProGlnPhelGlnIyrglySerIyValIyValIySHsProlaIaAspIlePro	80

Db	674	CTGTCCCCCAGAGTTCCAGTACGGCTCCAGAGTTACGTAGAGACCCCGGCACATCCCC	733		
Qy	81	AspTyrIylValLeuSerPheProGluGlyPheIylsTyrPgluIrgValMetAenPheGlu	100		
Db	734	GACTACAAAGAAAGCTGTCTCTCCCGAGGGCTTCAGATGGAGCGCGTGAAGAACTTCGAG	793		
Qy	101	AspGlyIglValIylValThrValIThrGlnIAspSerSerLeuGlnIAspGlyIylCysPheIleTyr	120		
Db	794	GACGGCGCGCTGTGTGACCGGTGACCCAGAGACTCTCCCTGTGAGGACGGCTCTTCATCTAC	853		
Qy	121	IylValIylAspPheIleGlyValIAspPheProSerAspGlyIProValIMetGlnIylsTyr	140		
Db	854	AAGGTGAAGTTCACTGGCGGTGAAGCTTCCCTCCGACGGCGCCCGTAAAGCAAGAAAGACT	913		
Qy	141	MetGlyIYrPgluIAsaSerThrGluIrgIleuTyrProIrgAspGlyValIleuIylGlu	160		
Db	914	ATGGGCTGTGGAGCCCTCCACCGAGCGCTGTATCCCGCGACGCGCGTGAAGGGCGAG	973		
Qy	161	IleHisIylValIleuIylsLeuIylsAspGlyIylHisTyrIleuValGluPheIylsSerIle	180		
Db	974	ATCCACAAAGGCGCTGAAGCTGAAGAGACGGGGCCACTACTGTGTGAGATTCAAGTTCATC	1033		
Qy	181	TyrMetAlaIylsIylsPProValGlnIleuProGlyIYrTyrTyrValIAspSerIylsLeuAsp	200		
Db	1034	TACATGTCCCAAGAAAGCCCGTGCAGCTGTCCCGGCTACTACTACGTGACTCAAGCTGCAC	1093		
Qy	201	IleThrSerHisAsnGluAspTyrTyrThrIleValGluGlnIYrGluIrgThrGluIylArg	220		
Db	1094	ATCAGCTCCCAACAGAGACTACACCATGTGTGAGACAGTACGAGCGCGCGAGGCGCG	1153		
Qy	221	HisHisIleuPheIleu 225			
Db	1154	CACCACTGTCTCTG 1168			
RESULT 62					
LOCUS	AJ851290	13079 bp	DNA		
DEFINITION	Cloning vector pRU1076 gusA gene for Beta-glucuronidase and deraedT4				
ACCESSION	AJ851290				
VERSION	AJ851290.1	GI:55724888			
KEYWORDS	beta-glucuronidase; Ds Red fluorescent protein T4; deraedT4 gene; gusA gene.				
SOURCE	Cloning vector pRU1076				
ORGANISM	Cloning vector pRU1076				
REFERENCE	1 other sequences; artificial sequences; vectors.				
AUTHORS	Karnunakaran, R. and Poole, P. S.				
TITLE	High throughput promoter probe vectors				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 13079)				
AUTHORS	Poole, P. S.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-OCT-2004) Poole P. S., School of AMS, University of				
FEATURES	Reading, Whiteknights, Reading, RG6 6AU, UNITED KINGDOM				
SOURCE	Location/Qualifiers				
	1. 13079				
	/organism="Cloning vector pRU1076"				
	/mol_type="other DNA"				
	/db_xref="taxon:299185"				
	/lab_host="Baccharichia coli"				
gene	186..863				
	/gene="deraedT4"				
CDS	186..863				
	/gene="deraedT4"				
	/codon_start=1				
	/transl_table=11				
	/product="Ds Red fluorescent protein T4"				
	/protein_id="CAH64919.1"				
	/db_xref="GI:55724889"				
	/translation="MASSSEIYKEMFIRFKYRMGSVNGHEFATIEGSGEGP PYRGTORA				
	KLVATVGGPLPFAWDILSFDQYGSKYTKYKHPADIDIDYKLKSLPBGFKRMVWNPBG				
	GVATVVDSSLDGCEPIYKVKFGVNPSPDGPWQKKTWGMEPSTERLYPRDGLKGE				



gene  
CDS  
IKKALIKDGGHYLVEFKSIYMAKKPVQLPGYYVDSKLDITSHNEDTIVEQYERAE  
GRHHLFL"  
885..2696  
/gene="gusa"  
985..2696  
/gene="gusa"  
/codon\_start=1  
/transl\_table=11  
/product="beta-glucuronidase"  
/protein\_id="CAH64920.1"  
/db\_xref="GI:55724890"

/translacion="WRRVEPTPTREIKKLDGMAFSLDRENGIDQRMWESALQESRA  
IAPSGFNDQPADADIRNYAGNVMYQREVFIPKGNAGRIYLRPAVTHYGTWNNQ  
EVMHQGGYTFEADVTPEYIAGSVRIIVCVNNELNQIIPGNAVITDENGKKQSY  
PHDFNVAIGHSVMLYTPNTMVDITVTHVADDCNHSADMQVANGVSVELRD  
ADQOVATGOGSGTLQOVNPHLMQGGYELCVTKASQTECDIPLARYGSRVAV  
KGEQFLTHKPPYPTFGFRHEDADIRGQFPNVLVMDHADMATGASVRSHPYA  
EEMLDWDEHGIIVIDETNAVQFNLSLIGPEAKPKKELISEBANGSETQAHICAI  
KELIARDKNHPSVVMVSIANEPTDRPQGAAREFALAEATRKLDTPREITCNWFC  
AHTDTISDLFDVLCLNRYGMYVQSGDLETAKVIEKELIAMQEKLHPITITEYGD  
TLAGHSMTDMMSEYOCAMLDMYHARVEDRVSAVVGEQVMNFADFATISQILRVGN  
KKGIFTRRKPSAFLIQKMTGNFGEKKPQGGKQ"

## ORIGIN

Alignment Scores:  
Pred. No.: 5,21e-108 Length: 13079  
Score: 1165.00 Matches: 216  
Percent Similarity: 97.33% Conservative: 3  
Best Local Similarity: 96.00% Mismatches: 6  
Query Match: 95.96% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AJ851290 (1-13079)

QY 1 MetAAGSerSerLyAsnValIleLySGluPhMeArArgPheLyValArgMetGluGly 20  
DB 186 ATGGCTCTCTCCGAGGACGTCATCAAGAGGTCATCGCTTCAAGGTGCGCATGAGGGCC 245  
QY 21 ThrValAsnGlyHISgluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 246 TCGGTAAAGCGCCACGATTCAGATCGAGGCGAGGCGGCGCCCTTACGAGGGC 305  
QY 41 HISenThrValLyLeuLyValThrLyGlyGlyProLeuProPheLatrpAspIle 60  
DB 306 ACCCAAGACCGCCAAAGCTGAAGGTGACCAAGGCGGCGCCCTTCCCTCGGAGCATC 365  
QY 61 LeuSerProGluPheGluTyrGlySerLyValTyrValIleHisProAlaAspIlePro 80  
DB 366 CTGTCCCCCAGTTCAGATCGGCTCCAGGTGTACGTGMAAGCACCCCCGGACATCCCC 425  
QY 81 AspTyrLyLyLeuSerPheProGluGlyPheLyValTrpGluArgValMetAsnPheGlu 100  
DB 426 GACTTCAAGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGGCGCTGATGAATCTCGAG 485  
QY 101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyr 120  
DB 486 GACGGGCGCTGTGACCGTGAACCGACCAAGACTCTCTCCCTGAGAGACGGCTGCTTCACTAC 545  
QY 121 LyValLyPheIleGlyValAlaAsnPheProSerAspGlyProValMetGluLyLeuThr 140  
DB 546 AAGGTGAAGTTCAATCGCGCTGAACCTTCCCTCGACGCGCCCGTAATGACGAAGAAGACT 605  
QY 141 MetGlyThrGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuLySGlyGlu 160  
DB 606 ATGAGCGGAGGACCTTCAACGAGCGCTGATACCCCGGACGAGCGGTGCTGAAGGGCGAG 665  
QY 161 ILeHisLyValLeuLyLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
DB 666 ATCCACAAAGCCCTGAAGCTGAAGGAGAGGGGCGCACTACCTGGTGAAGTTCAAGTTCATC 725  
QY 181 TyrMetAlaLyLyProValGluLeuProGlyTyrTyrTyrValAlaAspSerLyLeuAsp 200  
DB 726 TACATGGCCAAAGACCGGCTGACAGCTGCCGGCTACTACTACGTGACATCCAGGCTGAC 785

QY 201 ILeHisSerHisAsnGluAspArgTyrThrIleValGluGluTyrGluArgThrGluGlyArg 220  
DB 786 ATCACTCTCCCAACAGAGACTACACATCGTGAAGCATGACAGCGGCGCCGAGGGCCGC 845  
QY 221 HisHisLeuPheLeu 225  
DB 846 CACCACTGTTCTCTG 860

RESULT 63  
AF506025 681 bp DNA linear SYN 19-JUN-2002  
LOCUS AF506025  
DEFINITION Synthetic construct dimeric red fluorescent protein gene, complete  
ACCESSION AF506025  
VERSION AF506025.1 GI:21464833  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (24-Apr-2002) Pharmacology, UCSD, 9500 Gilman Drive, La Jolla, CA 92093, USA

## FEATURES

source  
1..681  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="from Discosoma sp.; Clontech vector pDRed-N1 with human optimized codon usage"  
1..681  
/note="dimer2; engineered variant of tetrameric red fluorescent protein from Discosoma deposited in Genbank  
Accession Number AF03369"  
/codon\_start=1  
/transl\_table=11  
/product="dimeric red fluorescent protein"  
/protein\_id="AAM54542.1"  
/db\_xref="GI:21464834"  
/translacion="WVASSEVDVIXEFMRFKVMESGVNGHEFEIGREGREPYEGTOT  
AKLTKTKGGRPLPRANDILISPORQVSKAYVHPADIPIYKKLSPFGFKRVMVPED  
GGVAVTVTQDSIQDGTLLIKVAFRGNTNPPDGPVWOKTMGWASSTERLYPRDGVLYG  
ETIHOALIKDKGGHYLVEFKSIYMAKKPVQLPGYYVDTKLDITSHNEDTIVEQYERAE  
GRHHLFL"

## ORIGIN

Alignment Scores:  
Pred. No.: 4,42e-105 Length: 681  
Score: 1121.00 Matches: 207  
Percent Similarity: 95.96% Conservative: 7  
Best Local Similarity: 92.83% Mismatches: 9  
Query Match: 92.34% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AF506025 (1-681)

QY 3 SerSerLyAsnValIleLySGluPhMeArArgPheLyValArgMetGluGlyThrVal 22  
DB 10 TCTCTCGAAGAGCTCATCAAGAGTTCATCGCGTTCAAGGTGCGATGAGGGCTCCGTC 69  
QY 23 AsnGlyHISgluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsn 42  
DB 70 AAGGCGCAAGATTCAAGATCGAGGCGAGGCGGCGCCCTTACGAGGCGACCCAG 129  
QY 43 ThrValLyLeuLyValThrLyGlyGlyProLeuProPheLatrpAspIleLeuSer 62

```
Db      130 ACCGCAAGCTGAAGTACCAAGGCGGCGCCCTTCCTGCTGGAGACATCTGTCC 189
Oy      63 ProGlnpHeGlnTYrGlySerIyValValIyVhIspRoAlaAspIleProAspTYr 82
Db      190 CCCGAGTTCAGAGTACGGCTCCCAAGGCGTACAGGAACACCCCGGACATCCCCGACTAC 249
Oy      83 LysIyLysLeuSerPhePheProGlnGlyPheIySTrpgIyuaRgVAlMetAAspPheGlnuAspGly 102
Db      250 AAGAAGCTGTCTCTTCCCGAGGCTTCAAGTGGAGCGCGTGAATACCTTCAGAGACGAC 309
Oy      103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTYrIyVal 122
Db      310 GCGCGTGAACCGTACCCAGAGACTCTCTCCCTGACAGACGGACCGCATGATCAAGGTG 369
Oy      123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIyLysThrMetGly 142
Db      370 AAGTTCGCGGACCACTTCCCGGACGGCGCCGCTAATGCAAGAAAGACATCGAGGC 429
Oy      143 TrpGlnAserThrGlnuArgLeuTYrProARgAspGlyValLeuIyGlyGlnuIleHis 162
Db      430 TGGAGGCTCTCACCGACGCGCTGTACCCCGGACGGCGTGTGAAGGCGGAGATCCAC 489
Oy      163 LysAlaLeuIyLysLeuYAspGlyGlyIyHleTYrLeuValGluPheIySerIleTYrMet 182
Db      490 CAGGCGCTGAAGCTGAAGAGACGGCGGCACTACCTGCTGGAGTTCAAGACCATCTACATG 549
Oy      183 AlaIyLysProValGlnLeuProGlyTYrTYrTYrValIAspSerLysLeuAspIleThr 202
Db      550 GCCAAGAGAGCCGCTGACACTGCGCGCTACTACTACGTGACACCAAGCTGGACATCCAC 609
Oy      203 SerHisangluAspTYrThrIleValGluGlnTYrGluArgThrGlnGlyAArgHis 222
Db      610 TCCCAACAAGAGACTACACATCTGGAAAGTACAGAGCGCTCCGAGGCGCGCACAC 669
Oy      223 LeuPheLeu 225
Db      670 CTGTTCTCTG 678

RESULT 64
AF506026 1395 bp DNA linear SYN 19-JUN-2002
LOCUS Synthetic construct tandem-dimer red fluorescent protein gene,
DEFINITION complete cds.
ACCESSION AF506026
VERSION AF506026.1 GI:21464835
KEYWORDS
SOURCE
ORGANISM synthetic construct
other sequences: artificial sequences.
REFERENCE
AUTHORS Campbell,R.E., Tour,O., Palmer,A.E., Steinhach,P.A., Baird,G.S.,
Zacharias,D.A. and Tsien,R.Y.
TITLE A monomeric red fluorescent protein
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (12), 7877-7882 (2002)
PUBMED 12060735
REFERENCE
AUTHORS Campbell,R.E. and Tsien,R.Y.
TITLE Direct Substitution
JOURNAL Submitted (24-Apr-2002) Pharmacology, UCSD, 9500 Gilman Drive, La
Jolla, CA 92093, USA
FEATURES
Source Location/Qualifiers
1..1395
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="from Diacosoma sp.; Clontech vector pBRed-N1 with
human optimized codon usage"
1..1395
/note="cdimer2(12); engineered variant of tetrameric red
fluorescent protein from Diacosoma deposited in GenBank
Accession Number AF03369"
/codon_start=1
```

```
/transl table=1
/product="tandem-dimer red fluorescent protein"
/protein_id="AA054543.1"
/db_xref="GI:21464836"
/translation="MVASSSDVYKEFNRFKVMGSGVNGHEFIEBGEGRPYEGTQT
AKLTKVKGGLPFPAMDILISPOFGSKAYVHPADIDYKLSPPBGRKMRVNFED
GGVTVVDDSLDGLTIYKVKFRGTNFPDGPVQVQKTMGMEASTELYTEODVKG
EIHQALFKDGGHLYVEFKTIYMAKKVQLPGYYVVDKIDITSHNEDYTVBOYERS
EGRHHLFLNGTSGTSGSSGTASREVIXEPMFKVMGSGVNGHEFIEBGEGRP
YEGQTAKLTKVKGGLPFPAMDILISPOFGSKAYVHPADIDYKLSPPBGRKMR
VMNPDGVVTVVDDSLDGLTIYKVKFRGTNFPDGPVQVQKTMGMEASTELYPR
DGVTKGEIHQALFKDGGHLYVEFKTIYMAKKVQLPGYYVVDKIDITSHNEDYTV
EYERSSEGRHHLFL"

ORIGIN

Alignment Scores:
Pred. No.: 1,06e-104 Length: 1395
Score: 1121.00 Matches: 207
Percent Similarity: 95.96% Conservative: 7
Best Local Similarity: 92.83% Mismatches: 9
Query Match: 92.34% Indels: 0
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AF506026 (1-1395)
Oy      3 SerSerLysAsnValIleIySGluPheMerArpPheLysValArgMetGluGlyThrVal 22
Db      10 TCTTCCGAGGACGTCATCAAAAGATTCAATGCGCTTCAAGGAGCGCATGAAGGCTCGTG 69
Oy      23 AsnGlyHisGluPheGlnIleGluGlyGluGlyArgProTYrGluGlyHisAsn 42
Db      70 AACGGCCACGAGTTCGAGATGAGGGGAGGGCGGCGCCCTACAGAGGACCCAG 129
Oy      43 ThrValIyLysLeuIyValThrIyGlyGlyProLeuProPheAlaIleAspIleLeuSer 62
Db      130 ACCGCAAGCTGAAGAGTGAACMAAGGGCGGCCCTGCGCTTGGCTGGACATCTGTGCC 189
Oy      63 ProGlnpHeGlnTYrGlySerIyValTYrValIyVhIspRoAlaAspIleProAspTYr 82
Db      190 CCCAGTTCACAGTACGGCTCCAGGCGGTACGTAAAGACCCCGCGACATCCCCGACTAC 249
Oy      83 LysIyLysLeuSerPheProGlnGlyPheIySTrpgIyuaRgVAlMetAAspPheGlnuAspGly 102
Db      250 AAGAAGCTGTCTCTTCCCGAGGCGTTCAGAGGAGCGCGTGAATACCTTCAGAGACGAC 309
Oy      103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTYrIyVal 122
Db      310 GCGCTGTGACCGTGACCCAGAGACTCTCTCCCTGACAGACGGACGCTGATCTACAGGTG 369
Oy      123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIyLysThrMetGly 142
Db      370 AAGTTCGCGGACCACTTCCCGGACGGCGCCGCTAATGCAAGAAAGACATCGAGGC 429
Oy      143 TrpGlnAserThrGlnuArgLeuTYrProARgAspGlyValLeuIyGlyGlnuIleHis 162
Db      430 TGGAGGCTCTCACCGACGCGCTGTACCCCGGACGGCGTGTGAAGGCGGAGATCCAC 489
Oy      163 LysAlaLeuIyLysLeuYAspGlyGlyIyHleTYrLeuValGluPheIySerIleTYrMet 182
Db      490 CAGGCGCTGAAGCTGAAGAGACGGCGGCACTACTGTGGAGTTCAAGACCATCTACATG 549
Oy      183 AlaIyLysProValGlnLeuProGlyTYrTYrTYrValIAspSerLysLeuAspIleThr 202
Db      550 GCCAAGAGAGCCGCGAGCGCTGCGCTACTACTACGTGACACCAAGCTGACATCCAC 609
Oy      203 SerHisangluAspTYrThrIleValGluGlnTYrGluArgThrGlnGlyAArgHis 222
Db      610 TCCCAACAAGAGACTACACATCTGGAAAGTACAGAGCGCTCCGAGGCGCGCACAC 669
Oy      223 LeuPheLeu 225
Db      670 CTGTTCTCTG 678
```

RESULT 65  
LOCUS AX824729 678 bp DNA linear PAT 11-DEC-2003  
DEFINITION Sequence 11 from Patent WO02068459.  
ACCESSION AX824729  
VERSION AX824729.1 GI:39750593  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. 678  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="hybrid coding sequence"  
ORIGIN  
Alignment Scores:  
Pred. No.: 7,03e-105 Length: 678  
Score: 1119.00 Matches: 206  
Percent Similarity: 96.89% Conservatave: 12  
Best Local Similarity: 91.56 Mismatches: 7  
Query Match: 92.17% Indels: 0  
Gaps: 0  
US-10-006-922a-12 (1-225) x AX824729 (1-678)  
QY 1 MetArgSerSerIlybAenValIlelybGluPhemeArGpHeleValArgMetGluGly 20  
DB 1 ATGAGCTGCAGCAAGACCTGATCAAGAGCTTCATCCGTTCAAGTCCGATGAGGCG 60  
QY 21 ThrValaenGlyHlaGluPhegiullegluGlyGluGlyValArgProTyrgluGly 40  
DB 61 ACCGTAAACGGCCACAGATTCCAGATCAAGGCGGAGGCGCGGCGCTTACGAGGCG 120  
QY 41 HlaenThValIlyleuValThrlVbGlyGlyProleuProPhelaItrPaPille 60  
DB 121 CACTGCAAGGCTGAAAGCTCATGTGACCAAGGCGGCGCTTCCCTTCGCTTCGACATC 180  
QY 61 LeuSerProGluPhegiullegluGlySerIlyValTyValIlybHlaProAlaAepIlePro 80  
DB 181 CTCAGGCCCCAGTTCCAGTACGCGACAGAGTGTACGTGAAGACACCCCGCCGACATCCCC 240  
QY 81 AApTyIlyleuValSerPheProGluGlyPheIyItrGluValMetAenPhegiul 100  
DB 241 GACTCAAGAGAGCTCAAGCTTCCCGAGGCGCTTCAAGTGGAGCGGATGAACTTCGAG 300  
QY 101 AApGlyIlyValIlyThrlValThrglnApsSerIleuGlnAepGlyCybPheIleTy 120  
DB 301 GACGCGGCGCTGTGTGACCGTGACCGACAGACAGCTTCAAGACGAGCTGCTTCATCTAC 360  
QY 121 LybValIyPheIleglyValAaenPheProSerAepGlyProValMetGlnIlybVthr 140  
DB 361 GAGGTAAAGTTCATCGCGCTGAACTTCCCAAGGACGCGCCCGTATGACGCGCGGAC 420  
QY 141 MetGlyItrGluIlyAseThrgluArgIleuTyProArgAepGlyValleuIySglIyGlu 160  
DB 421 CGGCGCTGGGAGGCGACGAGCGAGCGGCTTACCCCGGGAACGCGGCTCAAGGCGGAC 480  
QY 161 IlyHlaIlybAleuIyIlybAepGlyIlyHlaIlyleuValGluPhelySerIle 180  
DB 481 ATCCACATGGCCCTCCGCGCTCGAGGCGGCGCACCTCGTGAAGTTCAGAGATC 540  
QY 181 TyMetAlaIlybIyProValGlnIleuPProGlyTyTyTyTyValAaPserIlyleuA 200  
DB 541 TACATGGCCAAAGAGCCCGTGCAGCTCCCGGCTACTACTACGTGACACAGAGCTGCAC 600  
QY 201 IleThSerHlaaenGluAepTyThrlIleValGluGlnIlyGluArgThrgluGlyArg 220

DB 601 ATCACCACCCACACAGAGACTACACCATCTGTGACGATGACAGCGGACCGAGGCGCGG 660  
QY 221 HlaIlyleuPheleu 225  
DB 661 CACGACCTCTTCTCTC 675  
RESULT 66  
LOCUS AY678268 705 bp DNA linear SYN 17-DEC-2004  
DEFINITION Synthetic construct dimeric red fluorescent protein gene, complete cds.  
ACCESSION AY678268  
VERSION AY678268.1 GI:55420618  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. 705  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
1. 705  
/note="dimeric, engineered variant of dimeric red fluorescent protein dimer2 in Genbank Accession Number AF506025; N- and C-terminal replaced with equivalent residues from EGFP"  
/codon\_start=1  
/transl\_table=1  
/product="dimeric red fluorescent protein"  
/protein\_id="AAV52168.1"  
/translation="MWSKGEVYKEFMFKVMESGMNGHEFEIGEGRGREYEGTOTAKLVTYVQDSSLQDQTLILYKVMKRGNTNPPDPNWKMTGMEASTERLYPRDGLVKG EIHQALKDKDGGHYLVVEFKTLYMAKKPVQVLPETTYVDTKLDITSHNEDTIVEQYERS BERNHILFLYGMDELYK"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,08e-102 Length: 705  
Score: 1095.00 Matches: 201  
Percent Similarity: 94.57% Conservatave: 8  
Best Local Similarity: 90.95% Mismatches: 12  
Query Match: 90.20% Indels: 0  
Gaps: 0  
US-10-006-922a-12 (1-225) x AY678268 (1-705)  
QY 5 LybAenValIlelybGluPhemeArGpHeleValArgMetGluGlyThrlValaenGly 24  
DB 16 GAGAGAGTTCATAAAGAGTTCATGCGCTTCAAGGTGCGATGAGGCGCTCCATGAACGCG 75  
QY 25 HlaGluPhegiullegluGlyGluGlyIlyArgProTyrgluGlyHlaenThVal 44  
DB 76 CACGAGTTCAGATTCAGAGGCGGAGGCGGCGGCGGCGGCTTACGAGGCGACCCAGACCGCC 135  
QY 45 LybIlyleuValThrlVbGlyGlyProleuProPhelaItrPaPilleuSerProGln 64

```

Db      136 AAGCTGAAGGTGACCAAGGCGGCGCCCTTGCGCCCTTGCGCTGGGACATCTGTGCCCCCAG 195
Qy      65 PheGlnIYrGlySerIeValYrValIyShiSProAlaAepIleProAepTYrIySlyS 84
Db      196 TTCATGTACGGCTCCAAAGCGCTACGTGAAGCAACCCCGCGACATCCCGATTACAAAGAG 255
Qy      85 LeuSerPheProGluGlyPheIySTrpgIuArGValMeAenPheGluAapGlyVal 104
Db      256 CTGTCTCTTCCCGAAGGCTTCAGAGTGGAGCCCGCGATGAATCTTGAGAGACGGCGCTCTG 315
Qy      105 ValThrValIThrGlnAapSerSerIeGlnAapGlyCySpheIleTYrIyValIyAaphe 124
Db      316 GTGACCGGTGACCCAGACTCTCTCCCTGCGAGGCGGACCGCTGATCTACCAAGGTGAAGATG 375
Qy      125 IleGlyValAenPheProSerAapGlyProValMeGlnIySlyIyThMeGlyTYrpgIu 144
Db      376 CGCGGACCAACTTCCCCCGACGCGCCCGCTGAAGCAAGAAACACATGAGGTGGAG 435
Qy      145 AlaSerThrGluArGLeuTYrProArGAPGlyValIleuIySgIyGluIleHisIySAla 164
Db      436 GCCTTCACCGAGCGCTGTACCCCGCGACGCGCTGTGAGGCGGAGATCCACAGGCGC 495
Qy      165 LeuIySleuIyAapSgIyGlyYhiSTyIleuValIglupheIySerIeTYrMeAlaIyS 184
Db      496 CTGAAGCTGAAGGACGCGCGCCACTACCTGGTGAGATTCAAGACCATCTACATGGCCAA 555
Qy      185 LysProValGlnIleuProGlyTYrTYrTYrValIaapSerIySleuAapIleThSerHis 204
Db      556 AAGCCCGTGAACCTGCCGCTACTACTACGTGACCAAGCTGACATCATCTCCAC 615
Qy      205 AengIuAapTYrThIleValIgluGlnIyGluArGThnGluIyARgHishiIeuphe 224
Db      616 AACGAGACTACACATCTGTGAACGTHACAGCCCTCCGAGGCGGCCACCACTGTTC 675
Qy      225 Leu 225
Db      676 CTG 678

RESULT 67
LOCUS   AY678269 1431 bp DNA linear SYN 17-DEC-2004
DEFINITION Synthetic construct tandem-dimer red fluorescent protein gene, complete cds.
ACCESSION AY678269
VERSION  AY678269.1 GI:55420622
KEYWORDS
SOURCE  .
ORGANISM synthetic construct
          synthetic construct
          other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1431)
AUTHORS  Shaner,N.C., Campbell,R.E., Steinbach,P.A., Gilemans,B.N.,
          Palmer,A.E. and Tsien,R.Y.
TITLE    Improved monomeric red, orange and yellow fluorescent proteins
          derived from Discosoma sp. red fluorescent protein
JOURNAL  Nat. Biotechnol. 22 (12), 1567-1572 (2004)
PUBMED   15558047
REFERENCE 2 (bases 1 to 1431)
AUTHORS  Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and
          Tsien,R.Y.
TITLE    Direct Substitution
JOURNAL  Submitted (06-JUL-2004) Pharmacology, University of California, San
          Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA
FEATURES
         Source
         1..1431
         /organism="synthetic construct"
         /mol_type="other DNA"
         /db_xref="taxon:32630"
         1..1431
         /note="tdtomato; engineered variant of dimeric red
         fluorescent protein dimer2 in Genbank Accession Number
         AF506025; N- and C-termini replaced with equivalent
         residues from EGFP"
         /codon_start=1

```

```

/cirans1 table=11
/product="tandem-dimer red fluorescent protein"
/protein_id="AAV52169.1"
/db_xref="GI:55420623"
/translation="MVSKGSEVTKFMRFKVRMGSMNGHEFPIEGREGSPRYEGTOT
AKLKVTKGSLPFPAMDILSPQFMYSRAYVHPADIPDYKLSFPBGFPMKRVNMFED
GGATVTVQDSSLQGTLLIKVRKMGTFNPPDGPVMOKKTGMWMASTELIYPRDGLK
EIHQALTKDGGHYLVFPTIYMAKKVQVLPGYVYVTDIKDITSHNEDYTIYVOYBRS
EGRHHFLFHGTGTSGSSGSLASSENMAVLEKPMFKVRMGSMNGHEFPIEGREG
EGRPYEGTQTRAKLVTKGSLPFPAMDILSPQFMYSRAYVHPADIPDYKLSFPBGF
MKRVNMFEDGGATVTVQDSSLQGTLLIKVRKMGTFNPPDGPVMOKKTGMWMASTER
LYPRDGLVKGSHIQALTKDGGHYLVFPTIYMAKKVQVLPGYVYVTDIKDITSHNED
YTIYVOYBRSRHHFLFYGMDELK"

```

## ORIGIN

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 1095.00	1431	201	8	12	0	
Percent Similarity: 94.57%						
Best Local Similarity: 90.95%						
Query Match: 90.20%						
DB: 11						

US-10-006-922A-12 (1-225) x AY678269 (1-1431)

```

Qy      5  LysAenValIleIySgIupheMeArGpHeIyValArpMeGluGlyThrValAapGly 24
Db      16  GAGAGGTATCAAAAGATTCATCGCTTCMAAGTGGCATGAGGGCTCATAGACGGC 75
Qy      25  HisGluPheGluIleGluGlyGluIyArGProTYrGluGlyYhiAasnThrVal 44
Db      76  CACAGTTTCAGATCGAGGGGAGGGGAGGGCGCCCTTACAGAGGAGACCAGCCGCC 135
Qy      45  LysIeulysValIThrIySgIyIyProIeupProPheAlaTrpAapIleIeulysProGln 64
Db      136  AAGCTGAAGGTGACCAAGGCGCGCCCTGCTTCCTGCGGACATCTCTCCCGCCAG 195
Qy      65  PheGlnIYrGlySerIeValYrValIyShiSProAlaAepIleProAepTYrIySlyS 84
Db      196  TTCATGTACGGCTCCAAAGCGCTACGTGAAGCAACCCCGCGACATCCCGATTACAAAGAG 255
Qy      85  LeuSerPheProGluGlyPheIySTrpgIuArGValMeAenPheGluAapGlyVal 104
Db      256  CTGTCTCTTCCCGAAGGCTTCAGAGTGGAGCCCGCGATGAATCTTGAGAGACGGCGCTCTG 315
Qy      105  ValThrValIThrGlnAapSerSerIeGlnAapGlyCySpheIleTYrIyValIyAaphe 124
Db      316  GTGACCGGTGACCCAGACTCTCTCCCTGCGAGGCGGACCGCTGATCTACCAAGGTGAAGATG 375
Qy      125  IleGlyValAenPheProSerAapGlyProValMeGlnIySlyIyThMeGlyTYrpgIu 144
Db      376  CGCGGACCAACTTCCCCCGACGCGCCCGCTGAAGCAAGAAACACATGAGGTGGAG 435
Qy      145  AlaSerThrGluArGLeuTYrProArGAPGlyValIleuIySgIyGluIleHisIySAla 164
Db      436  GCCTTCACCGAGCGCTGTACCCCGCGACGCGCTGTGAGGCGGAGATCCACAGGCGC 495
Qy      165  LeuIySleuIyAapSgIyGlyYhiSTyIleuValIglupheIySerIeTYrMeAlaIyS 184
Db      496  CTGAAGCTGAAGGACGCGCGCCACTACCTGGTGAGATTCAAGACCATCTACATGGCCAA 555
Qy      185  LysProValGlnIleuProGlyTYrTYrTYrValIaapSerIySleuAapIleThSerHis 204
Db      556  AAGCCCGTGAACCTGCCGCTACTACTACGTGACCAAGCTGACATCATCTCCAC 615
Qy      205  AengIuAapTYrThIleValIgluGlnIyGluArGThnGluIyARgHishiIeuphe 224
Db      616  AACGAGACTACACATCTGTGAACGTHACAGCCCTCCGAGGCGGCCACCACTGTTC 675
Qy      225  Leu 225
Db      676  CTG 678

```

```

RESULT 68
LOCUS AF272711 876 bp mRNA linear INV 26-SEP-2000
DEFINITION Diacosoma sp. SSAL-2000 red fluorescent protein (FP593) mRNA,
complete cds.
ACCESSION AF272711.1 GI:10304306
VERSION AF272711.1
KEYWORDS Diacosoma sp. SSAL-2000
SOURCE Diacosoma sp. SSAL-2000
ORGANISM Diacosoma sp. SSAL-2000
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia;
Corallimorpharia; Discosomatidae; Discosoma.
REFERENCE 1 (bases 1 to 876)
Fradekov,A.F., Chen,Y., Ding,L., Barsova,E.V., Matz,M.V. and
Lukyanov,S.A.
Novel fluorescent protein from Diacosoma coral and its mutants
possesses a unique far-red fluorescence
FEBS Lett. 479 (3), 127-130 (2000)
JOURNAL 10981720
PUBMED
TITLE Novel fluorescent protein from Diacosoma coral and its mutants
posesses a unique far-red fluorescence
AUTHORS Fradekov,A.F., Chen,Y., Ding,L., Barsova,E.V., Matz,M.V. and
Lukyanov,S.A.
DIRECT SUBMISSION
JOURNAL Submitted (26-MAY-2000) Institute of Bioorganic Chemistry RAS,
Mikukho-Maklaya 16/10, Moscow 117871, Russia
FEATURES
SOURCE 1..876
/organism="Diacosoma sp. SSAL-2000"
/mol_type="mRNA"
/db_xref="taxon:137428"
1..876
/gene="FP593"
45..737
/gene="FP593"
/note="defP593"
/codon_start=1
/product="red fluorescent protein"
/protein_id="AAG16224.1"
/db_xref="GI:10304307"
/translation="MSSCKNVLIKEPMRPFVMEGTNGHEPEIKSGEGRPYEGHCSV
KLMVTKGGLPAPADILISQFQYSGSVYKHPADIPDYKLSPPSGPKERKRMPTEDG
GVATVSQDSLSKDGCEITVEKFTGVNFPDSGEPVQMRKRGMEASERLTPRGVTLKD
IHMLRLREGGHYLVFKSIYVKKRSVQLPGYYVDSKLDMTSHNEDYTVVEQYEKT
QGRHHPFKPLQ"
ORIGIN
Alignment Scores:
Pred. No.: 2,53e-101 Length: 876
Score: 1085.50 Matches: 199
Percent Similarity: 95.584 Conservative: 17
Best Local Similarity: 88.054 Mismatches: 9
Query Match: 89.424 Indels: 1
DB: 2 Gaps: 1
US-10-006-922A-12 (1-225) x AF272711 (1-876)
QY 1 McArGSeSerIySaAnValIlelySGluPhMeArGpHeIyVaIArgMeTGluGly 20
DB 45 ATGAGTTGTTCCAAAGATGATGATCAAGAGATCAAGGTTCAAGGTTGTTATGAAAGCA 104
QY 21 ThrValaSnGlyYHlsgluPhnegluilegluGlygluGlyValYargProTyrgluGly 40
DB 105 ACGGTCAATGGGCAAGAGTTGAAATTAAGGCAAGGTGAAGGAGGCTTACGAAGGT 164
QY 41 HlaaSnThrValaIySleuIySValThrlYsglyGlyProleuProPheIaIATpaaPile 60
DB 165 CACTGTTCCTTAAGCTTATGTTAAACCAAGGTGACCTTTCCTTTTGATATT 224
QY 61 leuSerProGlnhGlnhGlnTyrglySerIyVaIlyrValIyShIspRoAlaAaPILePro 80
DB 225 TTGTCAACCAATTTTCATGATGAAAGCAAGGTATATGTCAAAACCTTCCGCAATRCCA 284
QY 81 AspTyrlYelySleuSerPheProGluGlyPhelySITpGluArgValMetAaSnPheGlu 100

```

```

DB 285 GACTATATAAAGCTGTCTATTCTCGAGGATTAAATCGAAAGGTCATGAACCTTGAA 344
QY 101 AspGlyGlyValaIThrValaThrGlnaSpSerIeuGlnaSpGlyCyPheIleTy 120
DB 345 GACCGTGGCGTGGTACTGATCCCAAGATTCAGATTGAAAGCGCGCTGTTTCATCTAC 404
QY 121 LysValIySPhelIleGlyValaSnPheProSerAaPlyProValMetGlnIySlyThr 140
DB 405 GAGGTCAAGTTCATTTGGGGTGAACCTTCTGTATGAGCACTGTATACAGAGAGACA 464
QY 141 MetGlyTPGluAlaSerThrGluArgLeuTyProArGaSpGlyValIleuIySlyGlu 160
DB 465 CGGGCTGGAAAGCAAGCTCTGAAGCTTGTATCTCTGATGGGTGCTGAAGAAAGAGAC 524
QY 161 IleHlsYalAlaIySleuIySAspGlyGlyHlsTyrlYleuValGluPhelySerIle 180
DB 525 ATCCATATGGCTGTGAGGCTGGAAGAGCGGCCCATTTACCTGTTGAATTCAAAAGTATT 584
QY 181 TyrMetAlaIySlyPro---ValGlnleuProGlyTyTyTyTyValaAspSerIySleu 199
DB 585 TACATGTAAAGAACCTTCAGTGCAGTTCGCAAGCTACTATTATGTGACTCAAACTG 644
QY 200 AspIleThrSerHlsAsnGluAspTyThrIleValGluGlnTyrgluArgThrGluGly 219
DB 645 GATATGACGAGCCACCAAGATTAACACAGTCGTTGAGCACTATGAAAAAACCAAGGCA 704
QY 220 ArgHlsHlsleuPheleu 225
DB 705 CCCCAACATCCGTTCAATT 722
RESULT 69
LOCUS AX686894 876 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 17 from Patent WO0127150.
ACCESSION AX686894
VERSION AX686894.1 GI:29409472
KEYWORDS
SOURCE
ORGANISM Diacosoma sp.
Diacosoma sp.
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia;
Corallimorpharia; Discosomatidae; Discosoma.
REFERENCE 1
Lukyanov,S.A., Fradekov,A.F., Labae,Y.A., Matz,M.V. and Tersikh,A.
Anthozoa derived chromo/fluoroproteins and methods for using the
same
JOURNAL Patent: WO 0127150-A 17 19-APR-2001;
Clontech Laboratories Inc. (US)
FEATURES
source
1..876
/organism="Diacosoma sp."
/mol_type="unassigned DNA"
/db_xref="taxon:86600"
ORIGIN
Alignment Scores:
Pred. No.: 2,53e-101 Length: 876
Score: 1085.50 Matches: 199
Percent Similarity: 95.584 Conservative: 17
Best Local Similarity: 88.054 Mismatches: 9
Query Match: 89.424 Indels: 1
DB: 6 Gaps: 1
US-10-006-922A-12 (1-225) x AX686894 (1-876)
QY 1 McArGSeSerIySaAnValIlelySGluPhMeArGpHeIyVaIArgMeTGluGly 20
DB 45 ATGAGTTGTTCCAAAGATGATGATCAAGAGATCAAGGTTCAAGGTTGTTATGAAAGCA 104
QY 21 ThrValaSnGlyYHlsgluPhnegluilegluGlygluGlyValYargProTyrgluGly 40
DB 105 ACGGTCAATGGGCAAGAGTTGAAATTAAGGCAAGGTGAAGGAGGCTTACGAAGGT 164

```

```

QY      41 HisAenThrValIyLeuLyLeuValThrlYsgIyGlyProLeuProPheAlaTrpAspIle 60
      165 CACTGTTCGTAAGACTTATGTGTAAACAAGGTGACCTTTGGCCATTGCTTTGATATT 224
DB      61 LeuSerProGlnPheGlnTyrGlySerIyValIyValIyHisIProAlaAspIlePro 80
      225 TTGTACACACATTTTCAGTATGGAAGCAAGGATATGTCAACACCCGCGACATACCA 284
QY      81 AspTyrIyLysLeuSerPheProGlnIyPheIySTrpgIyLysIyValMetAsnPhgIu 100
      285 GACTATTAAGACCTTCATTCTCTGAGGATTTTAATGGAAAGGCTCATGAACTTTGAA 344
QY      101 AspGlyIyValIyValThrValThrGlnAspSerSerLeuGlnAspGlyCyAsPheIleTyr 120
      345 GACGGTGGCGTGTACTGTATCCCAAGTTCCATTGAAAGAAGCGCTGTTTCACTAC 404
DB      121 LysValIyPheIleGlyValIyAsnDheProSerAspGlyProValMetGlnIyLysThr 140
      405 GAGGTCAAGTTCATTGGGGTGAACCTTCTCTGATGACCTGTATGACAGAGAGACACA 464
QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIyGlu 160
      465 CGGGCTGGAAAGCCAGCTCTGAGCGTTGTATCTCGTATGGCGCTGTAAAGACAC 524
QY      161 IleHisIyAlaLeuIyLeuIyAspGlyIyHisIyTyrLeuValIyPheIySerIle 180
      525 ATCCATATGGCTCTGAGGCTGGAAGAGGCGGCCATTACCTCTGTAATTCAAAAGTATT 584
DB      181 TyrMetAlaIyLysPro--ValGlnLeuProGlyTyrTyrTyrValIyAspSerIyLeu 199
      585 TACATGCTAAAGAACCTTCAGTGCAGTTCGCCAGCTCTATTATGTGACTCCCAACCTG 644
QY      200 AspIleThrSerHisAsnGluAspTyrThrIleValIyGlnIyTyrGlnIyArgThrGlu 219
      645 GATATGACGAGCCACAAGAGATTACACAGTCGTTGAGCATGAAAGAACCAAGGGA 704
DB      220 ArgHisIleLeuPheIu 225
      705 CGCCACCATCGTTTCATT 722
QY
DB
RESULT 70
AF506027      678 bp      DNA      linear      SYN 19-JUN-2002
LOCUS      Synthetic construct monomeric red fluorescent protein gene,
DEFINITION      complete cds.
ACCESSION      AF506027.1 GI:21464837
VERSION      AF506027
KEYWORDS
SOURCE      .
ORGANISM      synthetic construct
              synthetic construct
              other sequences; artificial sequences.
REFERENCE      1 (bases 1 to 678)
AUTHORS      Campbell,R.E., Tour,O., Palmer,A.E., Steinbach,P.A., Baird,G.S.,
              Zacharias,D.A. and Tsien,R.Y.
TITLE      A monomeric red fluorescent protein
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 99 (12), 7877-7882 (2002)
PUBMED      12060735
REFERENCE      2 (bases 1 to 678)
AUTHORS      Campbell,R.E. and Tsien,R.Y.
TITLE      Direct Submission
JOURNAL      Submitted (24-Apr-2002) Pharmacology, UCSD, 9500 Gilman Drive, La
              Jolla, CA 92093, USA
FEATURES
SOURCE
1..678      Location/Qualifiers
              /organism="synthetic construct"
              /mol_type="genomic DNA"
              /db_xref="taxon:32630"
              /note="from Diacosoma sp.; Clontech vector pDsRed-N1 with
              human optimized codon usage"
              1..678
              /note="mrFP1; engineered variant of tetrameric red
              fluorescent protein from Diacosoma deposited in GenBank
              Accession Number AF03369"
CDS

```

```

/codon_start=1
/translation="monomeric red fluorescent protein"
/protein_id="AA05454.1"
/db_xref="GI:21464838"
/translation="MASEEDVIEKMFVKVRMGSGVNGHEPEIEGSGRPRPYGTQTA
KLKVTGGPLPFPAMDILSPQYGSKAYVHPALIPYLKLSPEGRKMEVNMPEEG
GVVTVTODSSLQDEPFYKYLKGTNPSGVPWQKKTWMEASTERYMPEDGALKE
IKMKLKLKDGHDVAEYKTTYMAKKPVQLPQAYKTDIKLDTSHNEDYTYIEQYERAE
GRHSTGA"
ORIGIN
Alignment Scores:
Pred. No.:      7,14e-95      Length:      678
Score:      1021.00      Matches:      192
Percent Similarity:      89.14%      Conservative:      5
Best Local Similarity:      86.88%      Mismatches:      24
Query Match:      84.10%      Indels:      0
DB:      11      Gaps:      0
US-10-006-922A-12 (1-225) x AF506027 (1-678)
QY      1 MecArgSerSerIyAsnValIleIyGluPheMetArgPheIyValArgMetGluIy 20
      1 ATGGCTCTCTCCGAGGACGTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGCG 60
DB      21 ThrValAsnGlyHisIyGluPheGluIleGluIyGluIyGluIyArgProTyrGluIy 40
      ::::
DB      61 TCCGTGAACGCGCACGAGTTCGAGATCGAGGCGAGGCGAGGCGCGCCCTTACAGAGGC 120
QY      41 HisAenThrValIyLeuLyLeuValThrlYsgIyGlyProLeuProPheAlaTrpAspIle 60
      121 ACCGAGCCGCGCAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTCGCTGGGACATC 180
DB      61 LeuSerProGlnPheGlnTyrGlySerIyValIyValIyHisIProAlaAspIlePro 80
QY      181 CTGTCCCTCAGTTCAGTACGCTCCAGGCTCTCAAGCATGAGCACCCCGCGACATCCCC 240
DB      81 AspTyrIyLysLeuSerPheProGluIyPheIySTrpgIyLysIyValMetAsnPhgIu 100
      241 GACTACTTGAAGCTGTCTCTCCCGAGGCTTCAAGTGGAGCGCGTATGTAACCTTGAG 300
QY      101 AspGlyIyValIyValThrValThrGlnAspSerSerLeuGlnAspGlyCyAsPheIleTyr 120
      301 GACGCGCGCGTGTGAGACGTGACCAAGACTCTCCCTGACAGACGCGAGTTCACTAC 360
DB      121 LysValIyPheIleGlyValIyAsnDheProSerAspGlyProValMetGlnIyLysThr 140
QY      361 AAGGTGAAGCTGCCGCGCACCAACTTCCCTCCAGCGGCCCGCTAATGCAAGAAAGACC 420
DB      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleIyGlu 160
      421 ATGGGCTGGAGGCTCTCACCGACGAGATGACCCGAGAGCGCGCTGAAAGGCGAG 480
QY      161 IleHisIyAlaLeuIyLeuIyAspGlyIyHisIyTyrLeuValIyPheIySerIle 180
      481 ATCAAGATGAGGCTGAAGCTGAAGAGCGCGGCCCATACAGCGCGCATGAACACACC 540
DB      181 TyrMetAlaIyLysProValGlnLeuProGlyTyrTyrTyrValIyAspSerIyLeuAsp 200
      541 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCGCTCAAGAACCGCATCAACTGAGAC 600
QY      201 IleThrSerHisAsnGluAspTyrThrIleValIyGlnIyTyrGlnIyArgThrGluIyArg 220
      601 ATCACTTCCCAACAGAGACTACACATCGTGAAGATGACGAGCGCGCGAGGCGCGC 660
DB      221 His 221
      661 CAC 663
QY
DB
RESULT 71
AB166761      696 bp      DNA      linear      SYN 02-OCT-2004
LOCUS      AB166761

```

DEFINITION Synthetic construct monomeric DsRed variant gene for red fluorescent protein, partial cds.  
ACCESSION AB166761  
VERSION AB166761.1 GI:53748885  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE  
1 Matsuo, K.  
Monomeric DsRed variant with plant codon usage  
Unpublished  
2 (bases 1 to 696)  
Matsuo, K.  
REFERENCE  
TITLE Direct Submission  
AUTHORS Submitted (08-MAR-2004) Ken Matsuo, RIKEN (The Institute of Physical and Chemical Research), Plant Science Center, 1-3-22  
TITLE URL: http://img.psc.riken.go.jp/strc/STRC20uab20index.htm,  
Tel.81-45-503-9575, Fax:81-45-503-9573)  
LOCATION/Qualifiers  
FEATURES  
source  
1..696  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
8..696  
/gene="monomeric DsRed variant"  
8..>696  
/gene="monomeric DsRed variant"  
/note="Synthetic monomeric DsRed protein with plant codon usage"  
/codon\_start=1  
/transl\_table=11  
/product="red fluorescent protein"  
/protein\_id="BAD52341.1"  
/db\_xref="GI:53748886"  
/translation="MASSEDVIREKFRPVKRGSGVNGHEFIEGEGEPYEGTQTA  
KLVTKGGLPPAMDILSPQFGSKAYVHPADIPDYKLSFPGFKRVMNEDG  
GVVTVQDSLDGSEIRYKLRGTNFPDGPVMOKTMGEASTRMYPEDGALGE  
IKRLKDKGSHYDAEVKTTMAKKVQLPGAVKTDIKDITSHNEDYTVVQYERAE  
GHSHTGARPSPES"  
ORIGIN  
Alignment Scores:  
Pred. No.: 7 37e-95 Length: 696  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922a-12 (1-225) x AB166761 (1-696)  
QY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20  
DB 8 ATGGCATCTTCCAGATGATGATTAAAGAGTTTATGATTCGAAGTGAAGATGAGGGA 67  
QY 21 ThrValAsnGlyYHieGluPheGluIleGluGlyValGluGlyArgProGlyArgGly 40  
DB 68 TCTGTTAAATGACATGATGATTCGAGATTGAGGGAAGAGAGGAAAGACCTTCAGAGGG 127  
QY 41 HLeAsnThrValLyLeuLyValThrLyGlyGlyProLeuProPheLeuATPAPile 60  
DB 128 ACTCAAGCTCTCAAGCTCAAGGTTACTAAGGGGAGACCATGCTTCGTTGGGATATC 187  
QY 61 LeuSerProGlnPheGlnTyArgLySerLyValTyValLySHiSPRoAlaapPile 80  
DB 188 CTGCTTCCACAAATTCAGATTCAGGCTTAAGCTTAACGTCACGACATCCAGCATATCCCA 247  
QY 81 AspTyLyLyLyLeuSerPheProGluGlyPheLySTPGLuArgValMetAnPheGlu 100  
DB 248 GACTATTGAAAGCTCTTCCAGAGGATTCAGGAGGAGATTATGAACTTCGAG 307

QY 101 AspGlyLyValValThrValThrGluAspSerLeuGlnAspGlyCyPheIleTyr 120  
DB 308 GATGAGGAGATTGTTACTGTTACTCAAGATTCCTCCCAAGAGGGAGATTCACTTAC 367  
QY 121 LysValLyPheIleGlyValAsnPheProSerAspGlyProValMetGlnLyLyThr 140  
DB 368 AAGGTGAAGCTCAGAGAACCTTCACTTCCACTGATGAGACCAAGTTATCAAAAGAACT 427  
QY 141 MetGlyTPGluAlaSerThrGluArgLeuTyProArgAspGlyValLeuLyGlyGlu 160  
DB 428 ATGGATGAGAGGCTTCTACTGAAAGATGATTCAGAAAGATGCTCTCCCAAGGAGAG 487  
QY 161 IleHisLyAlaLeuLyLeuLyAspGlyGlyHisTyTrLeuValGluPheLySerIle 180  
DB 488 ATCAAGATGAGACTCAAGCTGAAAGATGAGACATTAATGATGCTGAAGTCAAGACTACC 547  
QY 181 TyrMetAlaLyLyProValGlnLeuProGlyTyTyTyValAspSerLyLeuAsp 200  
DB 548 TACATGCTTACAGACCCAGTCCACTTCAGAGGCTTACAGACTGACATCAAGCTCGAT 607  
QY 201 IleThrSerHisAsnGluAspTyThrIleValGluGlnTyArgGluArgThrGluGlyArg 220  
DB 608 ATCACTCCCATTAACAGAGATTAACACATCGTGAACAAATAGAGAGGCTGAAGGTAGA 667  
QY 221 His 221  
DB 668 CAT 670  
RESULT 72  
DQ005472 4543 bp DNA circular SYN 07-MAY-2005  
LOCUS DQ005472  
DEFINITION Cloning vector pSAT6-RFP-N1, complete sequence.  
ACCESSION DQ005472  
VERSION DQ005472.1 GI:63002541  
KEYWORDS  
SOURCE Cloning vector pSAT6-RFP-N1  
ORGANISM Cloning vector pSAT6-RFP-N1  
REFERENCE  
1 (bases 1 to 4543)  
other sequences; artificial sequences; vectors.  
AUTHORS Chung, S.-M., Vyas, S., Kozlovskiy, S., Citovsky, V., and Tzfira, T.  
TITLE A versatile vector system for multiple gene expression in plants  
JOURNAL Unpublished  
2 (bases 1 to 4543)  
REFERENCES  
AUTHORS Chung, S.-M., Vyas, S., Kozlovskiy, S., Citovsky, V., and Tzfira, T.  
TITLE Direct Submission  
JOURNAL Submitted (12-APR-2005) Department of Biochemistry and Cell Biology, State University of New York at Stony Brook, Stony Brook, NY 11794, USA  
FEATURES  
source  
1..4543  
/organism="Cloning vector pSAT6-RFP-N1"  
/mol\_type="other DNA"  
/db\_xref="taxon:326515"  
441..762  
/note="CamV 35S promoter"  
768..1089  
/note="CamV 35S promoter"  
1190..1320  
/note="translational enhancer 5'-UTR from tobacco etch virus"  
1321..1379  
/note="multiple cloning site; MCS"  
1380..2057  
/codon\_start=1  
/product="RFP"  
/protein\_id="AAV25379.1"  
/db\_xref="GI:63002542"  
/translation="MASSEDVIREKFRPVKRGSGVNGHEFIEGEGEPYEGTQTA  
KLVTKGGLPPAMDILSPQFGSKAYVHPADIPDYKLSFPGFKRVMNEDG  
GVVTVQDSLDGSEIRYKLRGTNFPDGPVMOKTMGEASTRMYPEDGALGE  
IKRLKDKGSHYDAEVKTTMAKKVQLPGAVKTDIKDITSHNEDYTVVQYERAE  
GHSHTGARPSPES"  
QY 101 AspGlyLyValValThrValThrGluAspSerLeuGlnAspGlyCyPheIleTyr 120  
DB 308 GATGAGGAGATTGTTACTGTTACTCAAGATTCCTCCCAAGAGGGAGATTCACTTAC 367  
QY 121 LysValLyPheIleGlyValAsnPheProSerAspGlyProValMetGlnLyLyThr 140  
DB 368 AAGGTGAAGCTCAGAGAACCTTCACTTCCACTGATGAGACCAAGTTATCAAAAGAACT 427  
QY 141 MetGlyTPGluAlaSerThrGluArgLeuTyProArgAspGlyValLeuLyGlyGlu 160  
DB 428 ATGGATGAGAGGCTTCTACTGAAAGATGATTCAGAAAGATGCTCTCCCAAGGAGAG 487  
QY 161 IleHisLyAlaLeuLyLeuLyAspGlyGlyHisTyTrLeuValGluPheLySerIle 180  
DB 488 ATCAAGATGAGACTCAAGCTGAAAGATGAGACATTAATGATGCTGAAGTCAAGACTACC 547  
QY 181 TyrMetAlaLyLyProValGlnLeuProGlyTyTyTyValAspSerLyLeuAsp 200  
DB 548 TACATGCTTACAGACCCAGTCCACTTCAGAGGCTTACAGACTGACATCAAGCTCGAT 607  
QY 201 IleThrSerHisAsnGluAspTyThrIleValGluGlnTyArgGluArgThrGluGlyArg 220  
DB 608 ATCACTCCCATTAACAGAGATTAACACATCGTGAACAAATAGAGAGGCTGAAGGTAGA 667  
QY 221 His 221  
DB 668 CAT 670

terminator	2064	.2274	
terminator	/note="CAMV 35S terminator"		
CDS	complement (3483..4343)		
	/note="amp resistance"		
	/codon_start=1		
	/product="Bla"		
	/protein_id="AAV25360.1"		
	/db_xref="GI:63002543"		
	/translation="MSIQHFRVALIPFPAAFCLEVFAPHEFLVKVKAEDQLGARVGY		
	IEDLDNSGKILDSFRPRPPMSTFVLLCGAVALSIDGEOGLGRHYSONDLVE		
	YSPETEGHLDGDMQRELCSAAITMSMDLNTLTITGSKELTAALAHAMDVTLT		
	DRSPETLENBAIPDNERDITMPVAAATLTLRLKILTELLTLISROOLIIMVEADKVAQPL		
	LRSLAPGMFRLADSGAGERSRIIALALGPDGKPSRIIVITYTGSQATMDERRRQIA		
	EIGASLIIRKM"		
ORIGIN			
Alignment Scores:			
Pred. No.:	7,26e-94	Length:	4543
Score:	1021.00	Matches:	192
Percent Similarity:	89.14%	Conservative:	5
Best Local Similarity:	86.88%	Mismatches:	24
Query Match:	84.10%	Indels:	0
DB:	11	Gaps:	0
US-10-006-922A-12 (1-225) x DQ005472 (1-4543)			
QY	1	MeatGserSerlysaenValileYegluPheNetargPheLyValargMetGluY	20
Db	1380	ATGGCCCTCTCCGAGGAGCGTCATCAAGAGTTCAAGCGCTTCAAGGTGGCATGAGGCG	1439
QY	21	ThyValaIaenGlyHiegluPhegluIilegluYgluYgluYgluYargProTyrgluY	40
Db	1440	TTCGTGAACGCCACGAGTTTCGAATGAGGGCCAGGGCCGAGGGCCGCCCTCAAGAGGC	1499
QY	41	HisaenThrVallyleuLyValThyThreGlyglYProleuProPhealaTrpAspIle	60
Db	1500	ACCACAGCCGCCAGACTGAGAGTACCAAGGCGGCGCCCTCGCTTCGCTGGAGCATC	1559
QY	61	LeuSerProglInPheglInTyrglySerlyeValTyVallyshIaproalaAspIlePro	80
Db	1560	CTGTCCTCAGATTCCAGTACGCTCCACAGGCTTACGTAAGGACCCCGGCATCTCCC	1619
QY	81	AspTyrlYalyleuSerPheProgluYlyPheLystrPgluargValMetAsnPheglu	100
Db	1620	GACTACTTGAAGCTGCTCTCCCGAGGCGCTTCAAGTGGGAGCCGATGATCACTTCAG	1679
QY	101	AspGlyglYalyValThyralThrglInAspSerSerleuGlnaAspGlyCyPheIleTy	120
Db	1680	GACGCGGCGGTGTGACCGTGAACCAAGACTTCTCTCCGACGAGCGCGGAGTTATCTAC	1739
QY	121	LyValIlyPheIleGlyValaIasnPheProSerAspGlyProValMetGlnlyeThr	140
Db	1740	AAAGTGAAGCTGCGCGGACCAACTTCCCTCCGACGCGCCCGTAAATGAGAAAGACC	1799
QY	141	MeGlyYTrPgluIaSerThrgluargLeuTyTrProargAspGlyValleuLyeglyGlu	160
Db	1800	ATGGGCTGGAGGCGCTCCACCGAGCATGTACCCCAAGAGCGCGCGCTGMAAGGCGAG	1859
QY	161	IleHislyValaleuLyaleuLyasapGlyGlyHisTyrlleuValGluPheIlySerIle	180
Db	1860	ATCAAGATGAGGCTGAAGCTGAAGACGCGCGCACTACGACGCGGAGGTCAACACCC	1919
QY	181	TyMetAlaIyalylyProValGlnleuProglYTyTyTyTyValaIaasPserLyLeuAsp	200
Db	1920	TACATGGCCAAAGAACCCCTGCACTGCGCGGCCCTTCAAGACCGACATCAAGCTGAC	1979
QY	201	IleThrSerxiIaenGluAspTyTrThrlIeValGluGlnTyrgluargThrgluYarg	220
Db	1980	ATCACCTCCCAACAGAGACTACACCATGTGAAACAGTACGAGCGCGCGAGGCGCGC	2039
QY	221	His 221	
Db	2040	CAC 2042	

RESULT 73	
LOCUS DQ005474	4555 bp DNA circular SYN 07-MAY-2005
DEFINITION Cloning vector pSAT6-RFP-C1, complete sequence.	
ACCESSION DQ005474	
VERSION DQ005474.1	GI:63002546
KEYWORDS	
SOURCE	Cloning vector pSAT6-RFP-C1
ORGANISM	Cloning vector pSAT6-RFP-C1
REFERENCE	Other sequences; artificial sequences; vectors.
AUTHORS	1 (bases 1 to 4555)
TITLE	Chung,S.-M., Vyas,S., and Tzifira,T.
JOURNAL	A versatile vector system for multiple gene expression in plants
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 4555)
TITLE	Chung,S.-M., Vyas,S., Kozlovskiy,S., Citovsky,V. and Tzifira,T.
JOURNAL	Direct Submision
REFERENCE	Submitted (12-APR-2005) Department of Biochemistry and Cell
REFERENCE	Biology, State University of New York at Stony Brook, Stony Brook,
REFERENCE	NY 11794, USA
FEATURES	Location/Qualifiers
source	1..4555
	/organism="Cloning vector pSAT6-RFP-C1"
	/mol_type="other DNA"
	/db_xref="taxon:326517"
	441..762
promoter	/note="CaMV 35S promoter"
	768..1089
promoter	/note="CaMV 35S promoter"
	1190..1320
enhancer	/note="translational enhancer 5'-UTR from tobacco etch
	virus"
	1323..1397
CDS	/codon_start=1
	/product="RFP"
	/protein_id="AA25382.1"
	/db_xref="GI:63002547"
	/translation="MASSEDPVIEKFRMRFKVMEGSVNGHEFIEEGEGRPEYGTOTPA
	KVATVGGPILPFAMDDLSPOFGSKAYVGHADIDPYLKLSPFGKMRUNVNFEGQ
	GVTTVTDSSLDQGEPTLYKKLKGNTNPPSDGPMQCKTWKEASTERTYEDGALKEE
	IKRRLKLDGSHYDAEVKTYTMAKPPVQLFGAYKTIDKIDITSHNEDYITVEQYRAE
	GRSTGA"
	1998..2075
misc_feature	/note="multiple cloning site; MCS"
	2076..2286
terminator	/note="CaMV 35S terminator"
	complement (3495..4355)
CDS	/note="amp resistance"
	/codon_start=1
	/product="bla"
	/protein_id="AA25383.1"
	/db_xref="GI:63002548"
	/translation="MSIQHFRVALIIPFAFCILPFAHPETLVKVKDAEDQLGARVGY
	IELDLNSGKILVESPREERPMWSTFKVLQGVLRIDAGQOLGRIRHYSONDIVE
	YSLVTEKHLLDGETVRELCSAAITWSDNTLAAAILTTIGAPKEITLFLHMGDHYTDL
	DREPELNEAIIPDREDTTPVMAATTLRGLTGELTLASRQOLIDMEADKVAQPLD
	LRSLPAGWGIADKSGAGRGSGIITAAIGPDGKPRIVIVITYTSGSATMDENRROIA
	ETASLILKTM"
ORIGIN	
Alignment Scores:	
Pred. No.:	7.28e-94
Score:	1021.00
Percent Similarity:	89.14%
Best Local Similarity:	86.88%
Query Match:	84.10%
DB:	11
	Gaps: 0
US-10-006-922A-12 (1-225) x DQ005474 (1-4555)	
1 MetArgSerSerIysAsnValIleIleYgGluPheMetArgPheIysValaIa-gMetGluGly 20	



```

Db      1323 ATGGCCTCCTCCGAGAGCGTCATCAGAGATTCAATGCGCTTCAAGGTGCGCATGAGGCG 1382
Qy      21 ThrValaAnglyVHiegluPhlegluilegluylgluylgluylArgProTyrgluyl
Db      1383 TCCGTGAACGGCCACGAGTTCCAGATCCAGGGCGAGGGGCGGGCCGCCCTACAGAGGCG 1442
Qy      41 HisAntHrVallyleuLyValThrylysglyglYProleuProPheAlaTTPaspIle 60
Db      1443 ACCGAGCGCGCAAGCTGAGAGGTGACCAAGGGCGGGCCCTCCCTCCCTCGGAGCATTC 1502
Qy      61 LeuSerProgluInPhlegluIntyrglyserlyValTyrvallyVHisProAlaAspIlePro 80
Db      1503 CTGTCCCTCCAGATTCACAGTCCGCTCAAGGCGCTTACGTAAAGCACCCGCCACATCCCC 1562
Qy      81 AspTyLylyleuSerPheProgluylphelystrgluyluArgValMetAsnPhleglu 100
Db      1563 GACTACTGAACTGTCTTCCCTCCGAGGCTTCAAGTGGAGCGGTGATGAACTTCGAG 1622
Qy      101 AspGlyglYValValThryValThrglnAspSerSerLeuGlnAspGlyCyaspheileTy 120
Db      1623 GACGGCGGGGTGGTGAACGTGACCGAGACTCTCCCTCGAGAGCGGCGAGTTCAATCTAC 1682
Qy      121 LyvVallyspheileglyValAsnPheprosEterAspGlyProValMetGlnlylystr 140
Db      1683 AAGGTGAAGCTGCGCGGCAACAATTCCCTCCGAGCGGCCGTAAAGCAAGAAAGACC 1742
Qy      141 MetGlyTyrgluAlaSerThrgluArgleuTyProArgAspGlyValleuylsglylu 160
Db      1743 ATGGGTGGAGAGCTCCACCGAGCGAGTACCCCGAGAGCGGCGCCCTGAAGGGCGAG 1802
Qy      161 HislelyAlaleuLylyleuLyAspGlyglYHieTyrluValGluPhelyserIle 180
Db      1803 ATCAAGATGAGCTGAGAGCTGAGAGGAGCGGGGCCATACGACGCGAGGTCAAGACACC 1862
Qy      181 TyrMetAllylyleuProValGlnleuProglYTyTyTyValAspserlyleuAsp 200
Db      1863 TACATGCGCAAGAGCCCGGTGACGTGCGCGGCGCTTACAAAGCCAGCATCAAGCTGAGAC 1922
Qy      201 IleThSerHisAngluAspTyThryIleValGluGlnTyGluTyrgThrgluylArg 220
Db      1923 ATCACTCCCAACAAGAGCTACCATCTGTGAACATGACAGCGCGCGCGAGGCGCG 1982
Qy      221 His 221
Db      1983 CAC 1985

RESULT 74
LOCUS   DQ005475                4558 bp    DNA      circular SYN 07-MAY-2005
DEFINITION   Cloning vector pSAT6-RFP-N1, complete sequence.
ACCESSION   DQ005475
VERSION     DQ005475.1  GI:63002549
KEYWORDS
SOURCE
  ORGANISM   .
            Cloning vector pSAT6-RFP-N1
            Cloning vector pSAT6-RFP-N1
            other sequences: artificial sequences; vectors.
REFERENCE
  AUTHORS    Chung,S.-M., Vyasa,S., and Tzfira,T.
  TITLE      A versatile vector system for multiple gene expression in plants
  JOURNAL    Unpublished
  AUTHORS    2 (bases 1 to 4558)
  DIRECT SUBMISSION
  Submitted (12-APR-2005) Department of Biochemistry and Cell
  Biology, State University of New York at Stony Brook, Stony Brook,
  NY 11794, USA
FEATURES
  source
    1..4558
    /organism="Cloning vector pSAT6-RFP-N1"
    /mol_type="other DNA"
    /db_xref="taxon:326518"
    441..762
    /note="CamV 35S promoter"

```

```

promoter
  768..1089
  /note="CamV 35S promoter"
enhancer
  1190..1320
  /note="translational enhancer 5'-UTR from tobacco etch
  virus"
misc_feature
  1321..1394
  /note="multiple cloning site; MCS"
CDS
  1395..2072
  /codon_start=1
  /product="RFP"
  /protein_id="AAV25384.1"
  /db_xref="GI:63002550"
  /translation="MASESDVIKEFFMFKVMESGVNGHEPEIBGEGRGPEYGTQTA
  KLYTKGPGIPFADIIISPOFGSKYVKGPADIPDYIKLSPEGFKEFVNVNPEDE
  GVAVVTQDSISQDSERFIKYVLRGTPSPSDGPNWOKTMEASTERYPRDGALKEG
  IKMLKIKDGGHYDAEVKTTMAKKPVLPAYIKTDIKLDITSHNEDTIYEQYERAE
  GRHSYGA"
  2079..2289
  /note="CamV 35S terminator"
  complement(3498..4358)
  /note="amp resistance"
  /codon_start=1
  /product="Bla"
  /protein_id="AAV25385.1"
  /db_xref="GI:63002551"
  /translation="MSIOHPRVALIPPAACLPVFAHPETLYVKYKAEDOLGARVY
  IELDINSKILBSFPEERFPWMTFFKVLGCAVLSRTIDAGOBOLGRHIYSONDVE
  YSPVTEKHLDGMTRVRELCSAITWSDVTANLTLITGGPELTAFLHMGDHYTRL
  DRSEBELNEAIPNDERDTMPVAMATYRLKLTGELLTLASRQQLIDMEADKVAEPL
  LRSAIPAGMFIADKSGAGERGSGRIIAALGPDKPSRIIVITTSQATMDERNQIA
  EIGASLIRHM"

ORIGIN
Alignment Scores:
Pred. No.:      7.29e-94      Length:      4558
Score:          1021.00      Matches:      192
Percent Similarity:  89.14%      Conservative:  5
Best Local Similarity:  86.88%      Mismatches:   24
Query Match:      84.10%      Indels:      0
DB:              11          Gaps:          0

US-10-006-922A-12 (1-225) x DQ005475 (1-4558)
Qy      1 MetArgSerSerlyAsnValIlelysgluPhemeArgPheLyValArgMetGluyl 20
Db      1395 ATGGCCTCCTCCGAGAGCGTCATCAGAGATTCAATGCGCTTCAAGGTGCGCATGAGGCG 1454
Qy      21 ThrValaAnglyVHiegluPhlegluilegluylgluylgluylArgProTyrgluyl 40
Db      1455 TCCGTGAACGGCCACGAGTTCCAGATCCAGGGCGAGGGGCGGGCCGCCCTACAGAGGCG 1514
Qy      41 HisAntHrVallyleuLyValThrylysglyglYProleuProPheAlaTTPaspIle 60
Db      1515 ACCGAGCGCGCAAGCTGAGAGGTGACCAAGGGCGGGCCCTCCCTCCCTCGGAGCATTC 1574
Qy      61 LeuSerProgluInPhlegluIntyrglyserlyValTyrvallyVHisProAlaAspIlePro 80
Db      1575 CTGTCCCTCCAGATTCACAGTCCGCTCAAGGCGCTTACGTAAAGCACCCGCCGACATCCCC 1634
Qy      81 AspTyLylyleuSerPheProgluylphelystrgluyluArgValMetAsnPhleglu 100
Db      1635 GACTACTGAACTGTCTTCCCTCCGAGGCTTCAAGTGGAGCGGTGATGAACTTCGAG 1694
Qy      101 AspGlyglYValValThryValThrglnAspSerSerLeuGlnAspGlyCyaspheileTy 120
Db      1695 GACGGCGGGGTGGTGAACGTGACCGAGACTCTCCCTCGAGAGCGGCGAGTTCAATCTAC 1754
Qy      121 LyvVallyspheileglyValAsnPheprosEterAspGlyProValMetGlnlylystr 140
Db      1755 AAGGTGAAGCTGCGCGGCAACAATTCCCTCCGAGCGGCCGTAAAGCAAGAAAGACC 1814
Qy      141 MetGlyTyrgluAlaSerThrgluArgleuTyProArgAspGlyValleuylsglylu 160

```

Db	1815	ATGGGCTGGAGGGCTCCACCGAGCGATGATACCCCGAGCGAGCGGCGCTTGAAAGGGGAG	1874
Qy	161	HeHisylsalaleuylaleuylaleuylaspGlyGlyHisTyrLeuValGluPheIysSerIle	180
Db	1875	ATCAAGATGAGCGCTGAGCTGAGCTGAGGACGGCGGCCTACGACGCGGAGGTCAAGACCA	1934
Qy	161	TyrMetAlaValSerIleProValGluLeuProGlyTyrTyrTyrValAlaSerIleLeuAsp	200
Db	1935	TACATGGCGAAGAAAGCCCGTCAGCTGCGCGCGCTCAAGACCGACATCAAGCTGGAC	1994
Qy	201	IleHisSerHisAaenGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	1995	ATCACTCTCCCAACAGAGCACTACACCATCTGTGAACAGTACGAGCGCGCGAGGGCGGC	2054
Qy	221	His 221	
Db	2055	CAC 2057	
RESULT 75			
LOCUS	AJ851287	5311 bp	DNA
DEFINITION	Cloning vector pRU1144 mRFP1 gene for mono Red fluorescent protein..		
ACCESSION	AJ851287		
VERSION	AJ851287.1	GI:55724880	
KEYWORDS	mono Red fluorescent protein; mRFP1 gene.		
SOURCE	Cloning vector pRU1144		
ORGANISM	other sequences; artificial sequences; vectors.		
REFERENCE	1		
AUTHORS	Karnakakaran, R. and Poole, P.S.		
TITLE	High throughput promoter probe vectors		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 5311)		
AUTHORS	Poole, P.S.		
TITLE	Direct Submmission		
JOURNAL	Submitted (21-OCT-2004) Poole P.S., School of AMS, University of Reading, Whiteknights, Reading, RG6 6AJ, UNITED KINGDOM		
FEATURES			
source	1..5311		
	/organism="Cloning vector pRU1144"		
	/mol_type="other DNA"		
	/db_xref="taxon:299182"		
	/lab_host="Escherichia coli"		
gene	326..1003		
	/gene="mRFP1"		
CDS	326..1003		
	/gene="mRFP1"		
	/codon_start=1		
	/transl_table=1		
	/product="mono Red fluorescent protein"		
	/protein_id="CAH64892.1"		
	/db_xref="GI:55724881"		
	/translation="MASSEDIYKEFMRFKYMEGSVNGHEFIEBEGEGRPYEGTQTA		
	KLATVGGSLPFWMDILSPQFOYGSKAVKPADIPYLKLSPEGKMERVVMFEDGQK		
	GLVTVTDDSLDGEDEFTYKVLRGNTNPSPDQVWOKTMTGMEASTEMYPEDGLKKEE		
	IKMLKLKDGSHYAEVKTITMAKKPQVLPAAYITDILKDTISNEDYTIIVEQYERAE		
	GRHSTGA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	8,78e-94	Length:	5311
Score:	1021.00	Matches:	192
Percent Similarity:	89.14%	Conservative:	5
Best local Similarity:	86.88%	Mismatches:	24
Query Match:	84.10%	Indels:	0
DB:	11	Gaps:	0
US-10-006-922A-12 (1-225) x AJ851287 (1-5311)			
Qy	1	MetArgSerSerIysAaenValIleIysGluPheMetArgPheIysValArgMetGluGly	20
Db	326	ATGGCTCTCTCCGAGAGCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGGAAGGC	385

Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyValGluIleGluGlyAspProTyrGluGly	40
Db	386	TCCGTGAACGGCCACGAGTTTCGAGATGAGGGCGAGGGCCAGGGCCCTTACGAGGGC	445
Qy	41	HisAsnThrValIleuLysLeuLysValThrLysGlyLysProLeuProPheAlaTrpAspIle	60
Db	446	ACCAGACCGCCAGAGCTGAAGGTGACCAAGGGCGGCCCCCTGCGCTTGAGGACATC	505
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValIleHisProAlaAspIlePro	80
Db	506	CTGCCCCCTCAGTCCAGTACGGGTCCAGAGCCCTCACTGAAGCACCCCGCGCATCCCC	565
Qy	81	AspPheTyrLysLysLeuSerPheProGluGlyPheLysTrpGluLysValMetAsnPheGlu	100
Db	566	GACTACTTGAAGCTGCTCTCCCGAGGGCTTCAAGTGGAGGCCGTGATGACTTGGAG	625
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	626	GACGGCGGCGGTGTGACCGTGAACCGACGACTCCTCCTGGCAGAGCGCGAGTTGATGAC	685
Qy	121	LysValIleLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	686	AAGGTGAAGCTGGCGGCGCACCACTTCCCTCCGACGGCCCCGTAAAGCAGAAAGAACCC	745
Qy	141	MetGlyTyrTrpGluLysSerThrGluLysGluLeuTyrProArgAspGlyValLleuLysGly	160
Db	746	ATGGCGTGGAGAGGCTCTCACCGAGCGGATGACCCCGAGAGCGGCCCTGAAAGGGGAG	805
Qy	161	IleHisLysAlaLeuLysLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	806	ATCAAGATGAGGCTGGAAGCTGAAGAGAGGGGGCCACTACGACGCCGAGGTCAAGACACC	865
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	866	TACATGCGCAAGAGCCCGTGCAGCTCCCGGCGCTCAAGACCGACATCAACTGTGAC	925
Qy	201	IleHisSerHisAsnGluAspTyrThrIleValGluGlnTyrGluLysThrGluGlyArg	220
Db	926	ATCACCTTCCCAACAGAGGACTACACCATGTGTGAACAAGTACGAGCGCGCGAGGGCCGC	985
Qy	221	His 221	
Db	986	CAC 988	
RESULT 76			
AJ851291			
LOCUS	13079 bp	DNA	circular SYN 12-NOV-2004
DEFINITION	Cloning vector pRU161 gusA gene for beta-glucuronidase and mRFP1 gene for mono Red fluorescent protein.		
ACCESSION	AJ851291		
VERSION	AJ851291.1	GI:55724891	
KEYWORDS	beta-glucuronidase; gusA gene; mono Red fluorescent protein; mRFP1 gene.		
SOURCE	Cloning vector pRU161		
ORGANISM	Cloning vector pRU161		
REFERENCE	1	other sequences; artificial sequences; vectors.	
AUTHORS	Karnatakaran, R. and Poole, P.S.		
TITLE	High throughput promoter probe vectors		
JOURNAL	Unpublished		
REFERENCE	2	(bases 1 to 13079)	
AUTHORS	Poole, P.S.		
TITLE	Direct Submmission		
JOURNAL	Submitted (21-OCT-2004) Poole P.S., School of AMS, University of Reading, Whiteknights, Reading, RG6 6AD, UNITED KINGDOM		
FEATURES	Location/Qualifiers		
SOURCE	1..13079	/organism="Cloning vector pRU161"	
		/mol_type="other DNA"	
		/db_xref="taxon:299186"	
		/lab_host="Escherichia coli"	
		186..863	
		/gene="mRFP1"	
gene			

DB:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:
US-10-006-922A-12 (1-225) x AJ951291 (1-13079)	2.64e-93	1021.00	89.14%	86.88%
QY 1	1021.00	Matched: 192	Conservative: 5	Mismatches: 24
Db 186	ATGCGCTCCTCCGAGGACGTCATCAAGAGTTCAAGCCCTTCAAGGTGGCATGAGGGCC	245	Indels: 0	Gaps: 0
QY 21	ThrValAlaSerGlyHisLeuIleuGluPheMetCysPheLeuValArgMetGluGly	40		
Db 246	TCCGTGAACGGCCACGAGTTCGAGATTCGAGGGCGAGGGCGAGAGGGCCGCTTACAGAGGGC	305		
QY 41	HisAsnThrValLeuLeuLeuValThrIleuGlyGlyProLeuProPheAlaIlePro	60		
Db 306	ACCAGACCGCCAGACTGAGGTGACCAAGGGCGGCCCTTGGCTGGGACATC	365		
QY 61	LeuSerProGlnPheGlnIleGlySerIleValIleValLeuHisProAlaIlePro	80		
Db 366	CTGTCCTCCCTCAAGTTCAGTACGGCTCCAAAGGCTCAGTAAACACCCGCGACATCCCC	425		
QY 81	AspTyrIleValLeuSerPheProGluGlyPheIleGlyTyrGluValMetAsnPheGlu	100		
Db 426	GACTACTTGAAGCTCTCTCTCCCTCCAGAGGGCTTCAAGTGGAGCGGTGATGAATTCGAG	485		
QY 101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120		
Db 486	GACGGCGGGGTGTACCTGTGACCTGACCCAGACTCTCTCCCTGCAAGACGGCAAGTTCATCTAC	545		
QY 121	LeuValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnIleValThr	140		
Db 546	AAAGTGAAGCTGGCGGGGACCAACTTCCCTCCGACAGGGCCCGTAAATGACAGAAAGACC	605		
QY 141	MetGlyTyrProGlnAlaSerThrGluThrLeuTyrProValAspGlyValLeuIleuGlyGlu	160		

Db		606	ATGGCGCTGGAGGCGCTTCAACCGAGCGAATGTATCCCGGAGAGCGGCCCTTGAAAGGCGGAG	665
Oy	IleIleIeAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGlnPheLysSerIle	180		
Db		666	ATCAAATATGAGCGTTGAAGCTGAAGAAGACGGCGGCCTACAGACGCGGAGGTCAAGACCACC	725
Oy	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200		
Db		726	TTCATGTGGCAAGAAAGCCCGGTGCAGCTGCCCGGGCCTTACAAGACCATCAAGCTGAGAC	785
Oy	IleThrSerHisAsnGluAspTyrThrIleValGlnGlyIleTyrGlnArgThrGlnGlyArg	220		
Db		786	ATACCTCCCAACAGACGACTCACCATCTGTGAACATGACAGCGCGCCGAGGGCGGC	845
Oy	His 221			
Db	846 CAC 848			
RESULT 77				
AY678271		678 bp	DNA	linear SYN 17-DEC-2004
LOCUS	Synthetic construct monomeric yellow fluorescent protein gene,			
DEFINITION	complete cds.			
ACCESSION	AY678271			
VERSION	AY678271.1	GI:55420626		
KEYWORDS	.			
SOURCE	synthetic construct			
ORGANISM	synthetic construct			
REFERENCE	other sequences; artificial sequences.			
AUTHORS	1 (bases 1 to 678)			
TITLE	Shaner,N.C., Campbell,R.E., Steinbach,P.A., Glepmann,B.N., Palmer,A.E. and Tsien,R.Y. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein Nat. Biotechnol. 22 (12), 1567-1572 (2004) 15558047			
JOURNAL	2 (bases 1 to 678)			
PUBMED	Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and Tsien,R.Y. Direct Submmission Submitted 06-JUL-2004) Pharmacology, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA			
REFERENCE				
AUTHORS				
FEATURES				
source	Location/Qualifiers			
CDS	1..678			
	/organism="synthetic construct"			
	/mol_type="other DNA"			
	/db_xref="taxon:32630"			
	1..678			
	/note="mhoneydew, engineered variant of monomeric red fluorescent protein mRFP1 in Genbank Accession Number AF506027"			
	/codon_start=1			
	/transl_table=1			
	/product="monomeric yellow fluorescent protein"			
	/protein_id="YAYS217.1"			
	/db_xref="GI:55420627"			
	/translation="MASSSDVLIKEMFKFVKRMSSGVNGHEPEITRGSEGRPYESTQTR KLKVTGKGLPLPAWDILSPQFMWGSKAVYKRPADIPLYLKLSFEFGKRWKVMPEDGG KLVATVDSSLDGEFIYKYVLKGTNTPSPDGPWQKTGMVAATBEYIPEDGLKQEGE IKNRLLKTKDGHDGHDAEVKTTYVMKKPVQLPGAYKIDKDLDTSHNEDYTVIVEGYERAE GRHSTGA"			
ORIGIN				
Alignment Scores:				
Pred. No.:	1..92e-93	Length:	678	
Score:	1007.00	Matches:	188	
Percent Similarity:	88.69%	Conservative:	8	
Best Local Similarity:	85.07%	Mismatches:	25	
Query Match:	82.95%	Indels:	0	
DB:	11	Gaps:	0	

```
QY      1 MetAgsSerSerLyAsnValIleLySgluPhMeMetArSpheLyValArgMetGluGly 20
      1 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1 ATGGCTCTCTCCGAGAGACTCATCAAGAGATTCAATGCCCTTCAAGGTGGCATGAGAGGC 60
QY      21 ThrValaangLyHieGluPhneGluIleGluGlyGluGlyGluGlyArgProTyrgLugly 40
      21 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      61 TCCGGAACGGCCACGAGTTGAGATCGAGGGCGAGGGGCGAGGGCCGCCCTTACGAGGGC 120
QY      41 HisAsnThrValLyLeuLyValThrLySglYpProLeuProPheAlaTPAspIle 60
      41 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      121 ACCGAGACCGCCAAAGTGAAGTGAACCAAGGCGGCCCTTGCCTTGCCTGGAGCAATC 180
QY      61 LeuSerProGluInPhGluIleTyrgLySerLyValTyrgValLySHieProAlaAspIlePro 80
      61 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      181 CTGTCCCTCACTTCAATGTGGGGCTCCAGAGGCTTCAAGAGACACCCGCGCATATCCCC 240
QY      81 AspTyrgLySlyLeuSerPheProGluGlyPheLySTpGluArgValMetAsnPhneGlu 100
      81 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      241 GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGTGAATTTCCAG 300
QY      101 AspGlyGlyValIleValThrValThrGluAspSerSerLeuGluAspGlyCyPheIleTyrg 120
      101 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      301 GACGGCGCGGTGTGACCGTGAACGAGACTCTCTCCGCAAGAGCGCGAGTTCAATTCAC 360
QY      121 LySValLySPhelIleGlyValAsnPhPProSerAspGlyProValMetGluInLySlyThr 140
      121 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      361 AAGGTGAAGCTGGCGGCGCAACCACTTCCCTCCGAGCGGCCCTTAAATGCAAGAAAGACC 420
QY      141 MetGlyTyrgLualIleSerThrGluArgLeuTyrgProArgAspGlyValLeuLySglYglu 160
      141 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      421 ATGGCTGGGGCGGCACCAACGAGCGAGATGTACCCCGAGAGAGCGCGCTTGAAGGGCGAG 480
QY      161 IleHisLySAlaLeuLyLeuLyAspGlyGlyHieTyrgLeuValGluPhLySerIle 180
      161 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      481 ATCAAGATGAGGCTGAAGCTGAAGACGCGCGCACTACAGACCGCGAGATCAAGCAACC 540
QY      181 TyrMetAlaLySlyProValGluInLeuProGlyTyrgTyrgTyrgValAspSerLySleuAsp 200
      181 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      541 TACATGGCCAAAGACCCGTGCACCTGCCCGGCGCTTACAGATTGACGGAGCTGGAC 600
QY      201 IleThrSerHisAsnGluAspTyrgThrIleValGluGluInTyrgLuarGlyThrGluGlyArg 220
      201 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      601 ATCACTCCCAACAGAGACTACACCATGTGTGAACAGTACGAGCGCGGAGGGCGGC 660
QY      221 His 221
      221 |||
Db      661 CAC 663

RESULT 78
LOCUS   AY678270          678 bp    DNA          linear    SYN 17-DEC-2004
DEFINITION Synthetic construct monomeric orange-red fluorescent protein gene, complete cds.
ACCESSION AY678270
VERSION   AY678270.1 GI:55420624
KEYWORDS
SOURCE    .
ORGANISM  synthetic construct
          synthetic construct
REFERENCE 1 (bases 1 to 678)
          Shner,N.C., Campbell,R.E., Steinhach,P.A., Giepmans,B.N.,
          Palmer,A.E. and Tsien,R.Y.
          Improved monomeric red, orange and yellow fluorescent proteins
          derived from Drosophila sp. red fluorescent protein
          Nat. Biotechnol. 22 (12), 1567-1572 (2004)
PUBMED   1558047
TITLE     2 (bases 1 to 678)
          Shner,N.C., Campbell,R.E., Steinhach,P.A., Palmer,A.E. and
          Tsien,R.Y.
          Direct Submision
          Submitted (06-JUN-2004) Pharmacology, University of California, San
          Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA
```

```
FEATURES
    source
        location/Qualifiers
            1..678
            /organism="synthetic construct"
            /mol_type="other DNA"
            /db_xref="taxon:32630"
            1..678
            /note="Mtagerine; engineered variant of monomeric red
            fluorescent protein mRFP1 in Genbank Accession Number
            AF506027"
            /codon_start=1
            /transl_table=1
            /product="monomeric orange-red fluorescent protein"
            /protein_id="AAV52170.1"
            /db_xref="GI:55420625"
            /translation="MASSSDVIRKFKMFKVMEGVSNGHEPIRGGRRPEGNQTA
            KLRTRKGRPRPRMIDILSPQCYGSKAYVGHADIPDYIKLSPEGRKMRPVNMFQD
            GVATVDDSSLDGEFTYKVKLRGTNPSPGVPWQKTMGEASSERNYPEDGALKGE
            IKMLKIKDGDGHDAEVKTYTMKRPVQLPQAYKTDIKLDTSHNEDYTYLYERAE
            GRHSTGA"
    CDS
        ORIGIN
            Alignment Scores:
            Pred. No.:      6,22e-93      Length:      678
            Score:          1002.00      Matches:      189
            Percent Similarity: 88.24%      Conservative: 6
            Best Local Similarity: 85.52%      Mismatches: 26
            Query Match:      82.54%      Indels:      0
            DB:                11      Gaps:      0

US-10-006-922a-12 (1-225) x AY678270 (1-678)
QY      1 MetAgsSerSerLyAsnValIleLySgluPhMeMetArSpheLyValArgMetGluGly 20
      1 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1 ATGGCTCTCTCCGAGAGACTCATCAAGAGATTCAATGCCCTTCAAGGTGGCATGAGAGGC 60
QY      21 ThrValaangLyHieGluPhneGluIleGluGlyGluGlyGluGlyArgProTyrgLugly 40
      21 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      61 TCCGGAACGGCCACGAGTTGAGATCGAGGGCGAGGGGCGAGGGCCGCCCTTACGAGGGC 120
QY      41 HisAsnThrValLyLeuLyValThrLySglYpProLeuProPheAlaTPAspIle 60
      41 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      121 ACCGAGACCGCCAAAGTGAAGTGAACCAAGGCGGCCCTTGCCTTGCCTGGAGCAATC 180
QY      61 LeuSerProGluInPhGluIleTyrgLySerLyValTyrgValLySHieProAlaAspIlePro 80
      61 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      181 CTGTCCCTCACTTCAATGTGGGGCTCCAGAGGCTTCAAGGTGGAGCGCGCTTGAAGTTCCCC 240
QY      81 AspTyrgLySlyLeuSerPheProGluGlyPheLySTpGluArgValMetAsnPhneGlu 100
      81 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      241 GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGTGAATTTCCAG 300
QY      101 AspGlyGlyValIleValThrValThrGluAspSerSerLeuGluAspGlyCyPheIleTyrg 120
      101 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      301 GACGGCGCGGTGTGACCGTGAACGAGACTCTCTCCGCAAGAGCGCGAGTTCAATTCAC 360
QY      121 LySValLySPhelIleGlyValAsnPhPProSerAspGlyProValMetGluInLySlyThr 140
      121 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      361 AAGGTGAAGCTGGCGGCGCAACCACTTCCCTCCGAGCGGCCCTTAAATGCAAGAAAGACC 420
QY      141 MetGlyTyrgLualIleSerThrGluArgLeuTyrgProArgAspGlyValLeuLySglYglu 160
      141 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      421 ATGGCTGGGGCGGCCTCTCCAGAGGCTTCAAGTGGAGCGCGTGTGAATTTCCAG 480
QY      161 IleHisLySAlaLeuLyLeuLyAspGlyGlyHieTyrgLeuValGluPhLySerIle 180
      161 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      481 ATCAAGATGAGGCTGAAGCTGAAGACGCGCGCACTACAGACCGCGAGTCAAGCAACC 540
QY      181 TyrMetAlaLySlyProValGluInLeuProGlyTyrgTyrgTyrgValAspSerLySleuAsp 200
      181 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      541 TACATGGCCAAAGACCCGTGCACCTGCCCGGCGCTTACAGACGAGCATCAAGCTGGAC 600
QY      201 IleThrSerHisAsnGluAspTyrgThrIleValGluGluInTyrgLuarGlyThrGluGlyArg 220
      201 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```

Db      601 ATCACTCCACAAAGAGACTACCATCGTGAATTGTCAGACGCCGCCGAGGCCCC 660
Qy      221 His 221
Db      661 CAC 663

RESULT 79
LOCUS   AY679163                      702 bp    DNA     linear    SYN 21-NOV-2004
DEFINITION Synthetic construct mRFPmarrs-His6 gene, complete cds.
ACCESSION AY679163
VERSION   AY679163.1 GI:55793413
KEYWORDS
SOURCE    synthetic construct
          synthetic construct
          other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 702)
AUTHORS   Fischer,M., Haase,I., Simech,E., Gerisch,G. and
          Muller-Taubenberger,A.
TITLE      A brilliant monomeric red fluorescent protein to visualize
          cytoskeleton dynamics in Dictyostelium
JOURNAL    FEBS Lett. 577 (1-2), 227-232 (2004)
PUBMED    15527790
REFERENCE 2 (bases 1 to 702)
AUTHORS   Fischer,M.J. and Mueller-Taubenberger,A.
TITLE      Direct Submision
JOURNAL    Submitted (07-JUL-2004) Chemistry, Technical University Munich,
          Lichtenbergstr. 4, Garching 85747, Germany
FEATURES
     source
         1..702
             /organism="synthetic construct"
             /mol_type="other DNA"
             /db_xref="taxon:32630"
             1..702
             /note="red fluorescent protein"
     CDS
         /codon_start=1
         /transl_table=11
         /product="mRFPmarrs-His6"
         /protein_id="AAV65051.1"
         /db_xref="GI:55775541"
         /translation="MASSSDVYKEFMRFRVKNMGSYNGHEFPIEGREGPRVEGQTA
         KLVTKGGLPFPAMDILSPQFQSGKAYKHPADIPDYAKLFPBEGFKRMETNPEDG
         GVVTVQDSITLDGEFIYVKLGKTNPSDGPWOKTKGWELSTERMPEGLAKGE
         IKRLKLDKGHYDAEAKTYYKAKQVQLPGAYKTDIKDIIISHNEDYIVQYERAE
         GRHSTGAGSHHHHH"
         676..681
         /note="Region: linker"
         682..699
         /note="Region: histidine-tag; 6xHis"
ORIGIN
Alignment Scores:
Pred. No.:      4.26e-92      Length:      702
Score:          994.00        Matches:      186
Percent Similarity: 87.78%    Conservative: 8
Best Local Similarity: 84.16%  Mismatches: 27
Query Match:    81.88%       Indels:      0
DB:             11           Gaps:       0

US-10-006-922a-12 (1-225) x AY679163 (1-702)
Qy      1 MetArySerSerLyAsnValIleLyGluPhenMetArgPheLyValArgMetGluLy 20
Db      1 ATGGCATCATCAAGAAAGTGTATTAAAGATTATGAGATTCAAAAGTTAAATGGAAGT 60
Qy      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGlyGly 40
Db      61 TCAGTTAACGGTATGAAATTGAAATGGAAGTGAAGTGAAGTGAACATATGAAGT 120
Qy      41 HisAsnThrVallyLeuLyValThrLyGlyGlyProLeuProPheAlaTPAPAlle 60
Db      121 ACACAAACACGCAAAATTAAGTTACAAAGAGTGTCATTACATTGCGGATATT 180

```

```

Qy      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrVallyHisProAlaAspIlePro 80
Db      181 CTTCACCAACAATTTCAATATGTAGTAAGCAATATGTTAAATCCAGACGATATTTCCA 240
Qy      81 AspTyrLySlyLeuSerPheProGluGlyPheLyStrpGluArgValMetAsnPhGlu 100
Db      241 GATTATCTTAACTTTCAATTTCCAGAGGTTTAAATGGCAAAAGAAATTAAGAAATTTGAA 300
Qy      101 AspGlyLyValIleThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      301 GATGCGTGTTGTGAACAGTATCAAGATTCACATTAACAAGATGGGAATTTATTTAT 360
Qy      121 LyVallyAspPheIleGlyValAsnPhProSerAspGlyProValMetGlnLySlyThr 140
Db      361 AAGTTAAATTAAGAGGTACCAATTTTCCATCAAGATGTCACAGTAATCAAAAAAACT 420
Qy      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleLySlyGlu 160
Db      421 ATGGTTGGGAAGCATCAACAGAAAGATGTATCCAGAAAGTGTGCATTAAGGTGAA 480
Qy      161 IleHisLyAlaLeuLySlyLeuLyAspGlyGlyHisTyrLeuValGluPhLySeriLe 180
Db      481 ATTAATAATGAGATTAAACCTTAAGATGTGTGTCATTAATGATGCAGAACTTAACAAC 540
Qy      181 TyrMetAlaLySlySProValGlnLeuProGlyTyrTyrTyrValAspSerLySlyLeuAsp 200
Db      541 TATTAAGCTTAAAAACAAGTTCAATTTACAGGTGCTTTTAAAAACGATTAATTAATGAT 600
Qy      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      601 ATTATTTCACATTAAGAAGATTATATCTATTGTTGAACATATGAAGAAGCAGAAAGTGA 660
Qy      221 His 221
Db      661 CAC 663

RESULT 80
LOCUS   AY678265                      711 bp    DNA     linear    SYN 17-DEC-2004
DEFINITION Synthetic construct monomeric orange fluorescent protein gene,
          complete cds.
ACCESSION AY678265
VERSION   AY678265.1 GI:55420614
KEYWORDS
ORGANISM
          synthetic construct
          synthetic construct
          other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 711)
AUTHORS   Shaner,N.C., Campbell,R.E., Steinbach,P.A., Gilemans,B.N.,
          Palmer,A.E. and Tsien,R.Y.
TITLE      Improved monomeric red, orange and yellow fluorescent proteins
          derived from Discosoma sp. red fluorescent protein
JOURNAL    Nat. Biotechnol. 22 (12), 1567-1572 (2004)
PUBMED    15558047
REFERENCE 2 (bases 1 to 711)
AUTHORS   Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and
          Tsien,R.Y.
TITLE      Direct Submision
JOURNAL    Submitted (06-JUL-2004) Pharmacology, University of California, San
          Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA
FEATURES
     source
         1..711
             /organism="synthetic construct"
             /mol_type="Other DNA"
             /db_xref="taxon:32630"
         1..711
             /note="monomeric; engineered variant of monomeric red
             fluorescent protein mRFP1 in GenBank Accession Number
             AF506027; N- and C-terminal replaced with equivalent
             residues from EGFP"
             /codon_start=1
             /transl_table=11
             /product="monomeric orange fluorescent protein"

```

/protein\_id="AAV52165.1"  
/db\_xref="GI:55420615"  
/translation="MVSKGEEDNNALIKERPFKRVMEGVSNGHFEIEGEGRPYE  
GFQTRAKLVTKGGLPFPAMDILSPQFTGSAAYVHPADIPDYKLSPBGFKWERVM  
NFEEDGVVTVQDSLDQGEFIVKVKLAGTVPSPGPMQKMTGMEASSRMYPEDG  
ALKEGIKRRLKIKDGGHVTSEVKTYYAKKVPQLPQAYIVGIIKLDITSHNEDYIVGQ  
YERABGRHSTGMDLYK"

ORIGIN

Alignment Scores:  
Pred. No.: 1.4e-91 Length: 711  
Score: 989.00 Matches: 185  
Percent Similarity: 88.37% Conservative: 5  
Best Local Similarity: 86.05% Mismatches: 25  
Query Match: 81.47% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922a-12 (1-225) x AY678265 (1-711)

QY 7 ValIlelySGluPheMetArgPheIyValIArgMetGluGlyThrValIAsnGlyYHISglu 26  
:::|||||  
DB 34 ATCATCAAGAGATTATGAGCTTCAAGGTGGCATGAGAGGCTCCGTAAACGGCCACAG 93  
QY 27 PheGluIleGluGlyGluGlyValArgProTyrGluGlyYHISAsnThrValIySleu 46  
94 TTCGAGATCGAGGGGAGAGGGGCGGCGCCCTTCAAGGGCTTCAAGCGCTAAAGCTG 153  
QY 47 LysValIThrIySGlyIyProIeuProPheAlaTPRAspIleIeuSerProGlnPheGln 66  
154 AAGGTGACCAAGGGGTGGCCCTCGCTTCCGTGGACATCTCTCCCTCAGTTCCACC 213  
QY 67 TyrGlySerIyValIyValIyHISProAlaAspIleProAspTyrIyIySleuSer 86  
214 TACGGCTCAAGGGCTTACGTGAAGCAACCCCGCGACATCCCGACTACTTCAAGCTGTCC 273  
QY 87 PheProGluGlyPheIySGlyIyPheIyValIyMetAsnPheGluAspGlyIyValIyThr 106  
274 TTCGCCGAGGGCTTCAAGTGGAGCGCGCTGATGAATTCGAGAGCGGGCGCTGTGACC 333  
QY 107 ValIThrGlnAspSerSerIeuGlnAspGlyCybPheIleTyrIyValIyPheIleGly 126  
334 GTGACCCAGAGACTCTCCCTCGCAGAGCGCGGAGTTCACTCAAGGTGAAGCTGCGCGC 393  
QY 127 ValAsnPheProSerAspGlyProValIyMetGlnIyIySleuThrMetGlyTyrGluIyAser 146  
394 ACCAACTTCCCTCGAGAGCGCCCGTAAATGCAAGAAAGACCATGAGGCTGGAGGCTTCC 453  
QY 147 ThrGluIySGlyIyProIyAspGlyValIyIleuIySGlyIyIleIySleuIySleu 166  
454 TCCGAGCGGAGTATACCCGAGAGCGCGCGCTGAAGGGGAGATCAAGATGAGCTGAAG 513  
QY 167 LeuIyAspGlyIyYHISIyIyValIyGluPheIySerIleTyrMetAlaIyIySleuPro 186  
514 CTGAAGGAGAGCGGGCGACATCACTCCGAGGTCAAGACCACTTCAAGGCCCAAGAGGCC 573  
QY 187 ValGlnIeuProGlyIyTyrIyValIyAspSerIySleuAspIleIyThrSerHISAsnGlu 206  
574 GTGAGCTGCGCGGCGCTTACATCGTCGGCATCAAGTTGAGATCATCCCAACAGAG 633  
QY 207 AspTyrThrIleValIyGluGlnIyGluIyGluIyGluIyArgHIS 221  
634 GACTACACCATGTGTGAACAGTACGACGCGCGGCGCGCCAC 678  
DB

RESULT 81  
LOCUS AY678264 711 bp DNA linear SYN 17-DEC-2004  
DEFINITION Synthetic construct monomeric red fluorescent protein gene,  
complete cds.  
ACCESSION AY678264  
VERSION AY678264.1 GI:55420612  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct

other sequences; artificial sequences.

REFERENCE  
1 (bases 1 to 711)  
AUTHORS Shaner,N.C., Campbell,R.E., Steinbach,P.A., Giepmans,B.N.,  
Palmer,A.E. and Tsien,R.Y.  
TITLE Improved monomeric red, orange and yellow fluorescent proteins  
derived from *Drosophila* sp. red fluorescent protein  
Nac. Biotechnol. 22 (12), 1567-1572 (2004)  
JOURNAL 1558047  
PUBMED  
REFERENCE  
2 (bases 1 to 711)  
AUTHORS Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and  
Tsien,R.Y.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-2004) Pharmacology, University of California, San  
Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA  
FEATURES  
source  
1..711  
Location/Qualifiers  
1..711  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"

CDS

/note="mCherry; engineered variant of monomeric red  
fluorescent protein mRFP1 in Genbank Accession Number  
AF506027; N- and C-termini replaced with equivalent  
residues from EGFP"  
/codon\_start=1  
/product="monomeric red fluorescent protein"  
/protein\_id="AAV52164.1"  
/db\_xref="GI:55420613"  
/translation="MVSKGEEDNNALIKERPFKRVMEGVSNGHFEIEGEGRPYE  
GTQTRAKLVTKGGLPFPAMDILSPQFTGSAAYVHPADIPDYKLSPBGFKWERVM  
NFEEDGVVTVQDSLDQGEFIVKVKLAGTVPSPGPMQKMTGMEASSRMYPEDG  
ALKEIKORLKTGKHVDLVKTTYAKKVPQLPQAYIVGIIKLDITSHNEDYIVGQ  
YERABGRHSTGMDLYK"

ORIGIN

Alignment Scores:  
Pred. No.: 2.24e-91 Length: 711  
Score: 987.00 Matches: 184  
Percent Similarity: 88.84% Conservative: 7  
Best Local Similarity: 85.58% Mismatches: 24  
Query Match: 81.30% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922a-12 (1-225) x AY678264 (1-711)

QY 7 ValIlelySGluPheMetArgPheIyValIArgMetGluGlyThrValIAsnGlyYHISglu 26  
:::|||||  
DB 34 ATCATCAAGAGATTATGAGCTTCAAGGTGGCATGAGAGGCTCCGTAAACGGCCACAG 93  
QY 27 PheGluIleGluGlyGluGlyValArgProTyrGluGlyYHISAsnThrValIySleu 46  
94 TTCGAGATCGAGGGGAGAGGGGCGGCGCCCTTCAAGGGCTTCAAGCGCTAAAGCTG 153  
QY 47 LysValIThrIySGlyIyProIeuProPheAlaTPRAspIleIeuSerProGlnPheGln 66  
154 AAGGTGACCAAGGGGTGGCCCTCGCTTCCGTGGACATCTCTCCCTCAGTTCCACC 213  
QY 67 TyrGlySerIyValIyValIyHISProAlaAspIleProAspTyrIyIySleuSer 86  
214 TACGGCTCAAGGGCTTACGTGAAGCAACCCCGCGACATCCCGACTACTTGAAGCTGTCC 273  
QY 87 PheProGluGlyPheIySGlyIyPheIyValIyMetAsnPheGluAspGlyIyValIyThr 106  
274 TTCGCCGAGGGCTTCAAGTGGAGCGCGTGAATTCGAGAGCGGGCGCTGTGACC 333  
QY 107 ValIThrGlnAspSerSerIeuGlnAspGlyCybPheIleTyrIyValIyPheIleGly 126  
334 GTGACCCAGAGACTCTCCCTCGCAGAGCGCGAGTTCACTCAAGGTGAAGCTGCGCGC 393  
QY 127 ValAsnPheProSerAspGlyProValIyMetGlnIyIySleuThrMetGlyTyrGluIyAser 146  
394 ACCAACTTCCCTCGAGAGCGCCCGTAAATGCAAGAAAGACCATGAGGCTGGAGGCTTCC 453  
DB

QY 147 ThrGluArgLeuTyProArgAaspGlyValLeuLysGlyGluIleHisLeuAlaLeuLys 166  
:::|||||  
Db 454 TCCGAGCGGAGTATACCCGAGGAGCGGCCCTGAGAGGCGAGATCAAGCAGAGGCTGAAG 513  
QY 167 LeuLysAaspGlyGlyHisTyLeuValGluPheLysSerIleTyMetAlaLysLeuPro 186  
514 CTGAAGGAGCGGCGGCTACGACCGCTGAGTCAGACCACTTACAGGCCCAAGAGGCC 573  
QY 187 ValGluLeuProGlyTyTyTyTyValAaspSerLysLeuAaspIleThSerHisAanglu 206  
574 GTGAGCGTCCCGGCGCTTACAGCTCAACATCACTTGACATCACTCCCAACAG 633  
QY 207 AaspTyThrIleValGluGlnTyArgLysArgThrGluArgHis 221  
634 GACTACACCATGTGTGAAACATGACAGCGCGGCGGCGCCAC 678  
Db  
RESULT 82  
AY786536 681 bp mRNA linear INV 01-DEC-2004  
LOCUS AY786536  
DEFINITION Discosoma sp. LW-2004 fluorescent protein raspberry mRNA, complete cds  
ACCESSION AY786536  
VERSION AY786536.1 GI:55792806  
KEYWORDS  
SOURCE  
ORGANISM  
Discosoma sp. LW-2004  
Discosoma sp. LW-2004  
Eukaryota; Metazoa; Chordata; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
1 (bases 1 to 681)  
Wang, L., Jackson, W.C., Steinbach, P.A. and Tsien, R.Y.  
Evolution of new nonantibody proteins via iterative somatic  
hypermutation  
Proc. Natl. Acad. Sci. U.S.A. 101 (48), 16745-16749 (2004)  
JOURNAL  
PUBMED  
2 (bases 1 to 681)  
Wang, L., Jackson, W.C., Steinbach, P.A. and Tsien, R.Y.  
Direct Substitution  
Submitted (18-OCT-2004) Department of Pharmacology, University of  
California-San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA  
FEATURES  
source  
Location/Qualifiers  
1..681  
/organism="Discosoma sp. LW-2004"  
/mol\_type="mRNA"  
/db\_xref="taxon:301246"  
1..681  
/codon\_start=1  
/product="fluorescent protein raspberry"  
/protein\_id="AAV65486.1"  
/db\_xref="GI:55792807"  
/translation="MYSKGEVYKEFMRPKVMESGVNGHEFEIEGEGEPYEQTQT  
ATLKTKGSPRLPPAMITLSPOCMYSGKYVKKPADIDYLIKSPFGPKMERVMNPED  
GGVAVTVQSSLDGSEFTYKVKLRGTFPSDFVWQKTMGEASERYPRPDGALKG  
ENMKRLKDGGHYDAEVKTYWAKRPVLPGLKYTDIKDITSHEDYTIIVEYERA  
EGHSTYGA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,41e-91 Length: 681  
Score: 985.00 Matches: 185  
Percent Similarity: 88.02% Conservative: 6  
Best Local Similarity: 85.25% Mismatches: 26  
Query Match: 81.14% Indels: 0  
Gaps: 0  
Db: 2  
US-10-006-922a-12 (1-225) x AY786536 (1-681)  
QY 5 LysAaspValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAangly 24  
:::|||||  
Db 16 GAGAGAGTCTATCAAGAGTTCAATCCGCTTCAAGTGGCATGAGGCGCTCGGAACGCG 75  
QY 25 HisGluPheGluIleGluGlyGluGlyArgProTyArgLysHisAangThrVal 44  
|||||

Db 76 CACGAGTTCCGATCGAGGGCGAGGGCGGCCCTTACAGAGGCCACCAAGCGCC 135  
QY 45 LysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAapIleLeuSerProGln 64  
Db 136 AAGCTGAAGGTGACCAAGGGGTGGCCCTTGCTTGGCTGGGACATCTCTCCCTAG 195  
QY 65 PheGlnTyArgLysLeuValTyValLysHisProAlaAapIleProAaspTyLysLys 84  
Db 196 TGCATGTACGGCTCCAGAGGCTTACGTGAGACCCCGCCGACATCCCGCATCTTGAAG 255  
QY 85 LeuSerPheProGluGlyPheLysTrpGluArgValMetAanPheGluAaspGlyVal 104  
Db 256 CTGTCTCTCCCGAGGCGCTTCAAGTGGAGCGCGTGAATGAATTCAGAGCGCGCGTG 315  
QY 105 ValThrValThrGluAaspSerSerLeuGluAaspGlyCysPheIleTyLysValLysPhe 124  
Db 316 GTGACCGTGACCCGAGCATCTCTCTGAGAGCGCGGATTCATCTTCAAGGTGAAGCTG 375  
QY 125 IleGlyValAanPheProSerAaspGlyProValMetGlnLysLysThMetGlyTrpGlu 144  
Db 376 CGGGCACCACTTCTCCGACGCGCCGTATGCAAGAGACCAATGGGCTGGAG 435  
QY 145 AlaSerThrGluArgLeuTyProArgAaspGlyValLeuLysGlyGluIleHisLysAla 164  
Db 436 GCCTCTCCGAGCGGATGTACCCCGAGAGCGCGCCCTGAAAGGCGGAGATGAATGAGG 495  
QY 165 LeuLysLeuLysAaspGlyGlyHisTyLeuValGluPheLysSerIleTyMetAlaLys 184  
Db 496 CTGAAGCTGAGAGAGCGGCGCCATACGACCGCGGTCAAGACCACTTACATGGCCAAAG 555  
QY 185 LysProValGlnLeuProGlyTyTyTyTyValAaspSerLysLeuAaspIleThSerHis 204  
Db 556 AAGCCGTGAGCTCCCGCGCCCTTACAGACCGGACATCAAGCTGACCTCCAC 615  
QY 205 AangLysAaspTyThrIleValGluGlnTyArgLysArgThrGluArgHis 221  
Db 616 AACGAGCTACACATGTGTGAAACATGACAGCGCGCGGAGCGCGCAC 666  
Db  
RESULT 83  
AY678266 711 bp DNA linear SYN 17-DEC-2004  
LOCUS AY678266  
DEFINITION Synthetic construct monomeric red-orange fluorescent protein gene, complete cds.  
ACCESSION AY678266  
VERSION AY678266.1 GI:55420616  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
other sequences; artificial sequences.  
1 (bases 1 to 711)  
Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.,  
Palmer, A.E. and Tsien, R.Y.  
Improved monomeric red, orange and yellow fluorescent proteins  
derived from Discosoma sp. red fluorescent protein  
Nat. Biotechnol. 22 (12), 1567-1572 (2004)  
JOURNAL  
PUBMED  
2 (bases 1 to 711)  
1558047  
Shaner, N.C., Campbell, R.E., Steinbach, P.A., Palmer, A.E. and  
Tsien, R.Y.  
Direct Substitution  
Submitted (06-JUL-2004) Pharmacology, University of California, San  
Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA  
FEATURES  
source  
Location/Qualifiers  
1..711  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
1..711  
/note="msRaspberry; engineered variant of monomeric red  
fluorescent protein mRFP1 in Genbank Accession Number  
AF506027; N- and C-terminal replaced with equivalent  
residues from EGFP"  
/codon\_start=1  
CDS  
11

```

/transl_table=11
/product="monomeric red-orange fluorescent protein"
/protein_id="AAV52166.1"
/db_xref="GI:55420617"
/translation="MVSKEENNMNAIIEFMRFKVRMGSVNGHEFIEGEGEPYE
GTOTAKLKYTKGGLPFPAMDILTPFTYGSKAYVGPADIPYLTLSPEEGKMERVM
NPEGDGVTVTODSLDGEFITYKVKRGTPSPGPMOKTGMWASSEMYPEDG
ALKEIKMRILKDKDGHYDAEVKITYAKKPVQLPGATYIVGIKDITSHNEDYITVEL
YERAEGRNSTGMDRLYK"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 6,03e-90 Length: 711
Score: 973.00 Matches: 182
Percent Similarity: 87.44% Conservative: 6
Best Local Similarity: 84.65% Mismatches: 27
Query Match: 80.15% Indels: 0
DB: 11 Gaps: 0
US-10-006-922a-12 (1-225) x AYT678266 (1-711)

```

```

OY 7 ValIleLyGluPhMeLArGPhLeYValArGMeLGlUgLYThrValAsnGlyHisGlu 26
:::|||||
DB 34 ATCATCAAGAGATTGATCGCTTCAAGGTGCGCATGAGGCGCTCGTGAACGGCCACAG 93
OY 27 PheGluIleGluGlyGluGlyValArgProTYrGluGlyHisAsnThrValysLeu 46
94 TTCGAGATCGAAGGCGAGGCGCGCGCCCTTACAGAGGCGCACCGACCGCCACAGCTG 153
OY 47 LysValThrLyGlyGlyProLeuProPheAlaTPAerPLeuSerProGlnPheGln 66
154 AAGGAGACCAAGGATGGCCCTTGCCTTGGCATCTCAACCCCACTTCAACC 213
OY 67 TYrGlySerLySValTYrValLyHisProAlaAerPLeuSerTYrLyLeuSer 86
214 TACGGCTCCAAAGCTTACGTGAAGCACCCCGCGCATCCCGACTTGAAGCTGTCC 273
DB 87 PheProGluGlyPheLySerTrpGluArgValMetAsnPheGluAspGlyGlyValValThr 106
274 TTCGGAGGGCTTCAAGTGGAGGCGCGTGAATGATCTTGAAGAGCGCGCGTGGTGAAC 333
OY 107 ValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTYrLySValLyPheIleGly 126
334 GTGAGCCAGGAGTCTCTCTCGAGAGCGCGCATTCATCTACAAGTGAAGCTGGCGGC 393
DB 127 ValAsnPheProSerAspGlyProValMetGlnLySlySerThrMetGlyTYrGluAlaSer 146
394 ACCAATTCCCTCCGACGCGCCCGTAATGACAGAAAGACCATGGGCTGGAGGCTGCC 453
OY 147 ThrGluArgLeuTYrProArgAspGlyValLeuLySlyGlyGluIleHisValAlaLeu 166
454 TCCGAGCGGATTAACCCGAGAGCGCGCGCTGAAGGGGAGATCAAGATGAGGCTGAAG 513
DB 167 LeuLyAspGlyGlyHisTYrLeuValGluPheLySerIleTYrMetAlaLySlyPro 186
514 CTGAAGGAGACGGGGCCACTACAGCGCTGAGGTCAAGACCACTACAAGGCCAAGAGCCC 573
OY 187 ValGlnLeuProGlyTYrTYrTYrValAspSerLySlyAspPLeuSerHisAsnGlu 206
574 GTGAGAGTCCCGCGCGCTTACATCGTCGCGCATCAAGTTGAGCATCACTCCCAACAGAG 633
DB 207 AspTYrThrIleValGluGlnTYrGluArgThrGluGluArgHis 221
634 GACTACACCATGTGAACTGTACGAACGCGCGAGGGCGGCCAC 678

```

```

RESULT 84
AY786537 681 bp mRNA linear INV 01-DEC-2004
DEFINITION Diacosoma sp. LM-2004 fluorescent protein plun mRNA, complete cds.
ACCESSION AY786537
VERSION AY786537.1 GI:55792808
KEYWORDS
SOURCE Diacosoma sp. LM-2004

```

## ORGANISM

Diacosoma sp. LM-2004

Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Diacosomatidae; Diacosoma.

## REFERENCE

1 (bases 1 to 681)

## AUTHORS

Wang, L., Jackson, W.C., Steinbach, P.A. and Tsien, R.Y.

## TITLE

Evolution of new nonantibody proteins via iterative somatic hypermutation

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 101 (48), 16745-16749 (2004)

## PIRMBED

15556995

## REFERENCE

2 (bases 1 to 681)

## AUTHORS

Wang, L., Jackson, W.C., Steinbach, P.A. and Tsien, R.Y.

## TITLE

Direct Substitution

## JOURNAL

Submitted (18-OCT-2004) Department of Pharmacology, University of California-San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

## FEATURES

Location/Qualifiers

## SOURCE

1..681

## CDS

1..681

## ORIGIN

```

Alignment Scores:
Pred. No.: 7,24e-90 Length: 681
Score: 972.00 Matches: 181
Percent Similarity: 87.10% Conservative: 8
Best Local Similarity: 83.41% Mismatches: 28
Query Match: 80.07% Indels: 0
DB: 2 Gaps: 0
US-10-006-922a-12 (1-225) x AY786537 (1-681)

```

```

OY 5 LysAsnValIleLyGluPhMeLArGPhLeYValArGMeLGlUgLYThrValAsnGly 24
:::|||||
DB 16 GAGGAGGTCATCAAGAGATTGATCGCTTCAAGAGACATGAGGCGCTCGTGAACGGC 75
OY 25 HisGluPheGluIleGlyGlyGlyValArgProTYrGluGlyHisAsnThrVal 44
76 CACGAGTTCGAGATCGAAGGCGAGGCGCGCGCCCTTACAGAGGCGCACCGACCGCC 135
DB 45 LysLeuLySValThrLyGlyGlyProLeuProPheAlaTPAerPLeuSerProGln 64
136 AAGCTGAAGGTGACCAAGAGTGGCCCTTGCCTTGGCACTTCTCCCTCAG 195
OY 65 PheGlnTYrGlySerLySValTYrValLyHisProAlaAerPLeuSerTYrLyLeu 84
196 ATCATGTAACGGCTTCAAGGCTTACGTGAAGCACCCCGCGCATCCCGCATCTTGAAG 255
DB 85 LeuSerPheProGluGlyPheLySerTrpGluArgValMetAsnPheGluAspGlyVal 104
256 CTGTCTTCCCGCGAGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAGAGCGGCGCTG 315
OY 105 ValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTYrLySValLyPhe 124
316 GTGACCTGACCCAGGAGTCTCTCTTGAAGAGCGCGCATTCATCTACAAGTGAAGTGTG 375
DB 125 IleGlyValAsnPheProSerAspGlyProValMetGlnLySlySerThrMetGlyTYrGlu 144
376 CGCGGACCAACTTCCCTCGAGCGCGCGTGAATGAGAGACCACTAGGCGTGGAG 435
OY 145 AlaSerThrGluArgLeuTYrProArgAspGlyValLeuLySlyGlyGluIleHisValAla 164
436 GCTTCTCCGAGCGGAGTGAACCCGAGAGCGCGCCCTTGAAGGGCGAGATGAAGTGAAG 495

```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 12, 2006, 11:20:58 ; Search time 478 Seconds  
(without alignments)  
3137.146 Million cell updates/sec

Title: US-10-006-922a-12  
Perfect score: 1214  
Sequence: 1 MRSKRVKIFKFRFKRMKG.....EDYIVGEYERTEGRHFL 225

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4996997 seqs, 3332346308 residues  
Total number of hits satisfying chosen parameters: 83

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DRV=xlp  
-Q=cg21\_1/USFTO.spool\_p/US10006922/runat\_10012006.162503.19106/app\_query.fasta\_1.391  
-DB=N.Geneseq -QMT=faetap -SUFFI=p2nminscr.rng -MINMATCH=0.1 -LOOFCLE=0  
-LIST=500 -DICALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=69 -ALIGN=500  
-MODE=LOCAL -OUTMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922 @CGN 1.1 1096 @runat\_10012006.162503.19106 -NCPU=6 -ICPU=3  
-NO MAP -LARGESUBERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N.Geneseq 21: \*  
1: geneseqn19808: \*  
2: geneseqn19908: \*  
3: geneseqn20008: \*  
4: geneseqn20018: \*  
5: geneseqn20028: \*  
6: geneseqn20038: \*  
7: geneseqn20048: \*  
8: geneseqn20058: \*  
9: geneseqn20068: \*  
10: geneseqn20078: \*  
11: geneseqn20088: \*  
12: geneseqn20098: \*  
13: geneseqn20108: \*  
14: geneseqn20118: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	100.0	678	4	AAH47654
2	1214	100.0	678	4	AAAD13053
3	1214	100.0	678	4	AAAD11142
4	1214	100.0	678	6	AAAG95905

5	1214	100.0	678	6	AAAG95922
6	1214	100.0	678	6	AAAG95921
7	1214	100.0	678	6	AAAG95920
8	1214	100.0	678	6	AAAG95919
9	1214	100.0	678	6	AAAG95918
10	1214	100.0	678	10	AAAG95917
11	1214	100.0	678	10	AAAG95916
12	1214	100.0	678	11	AAAG95915
13	1214	100.0	678	11	AAAG95914
14	1214	100.0	678	12	AAAG95913
15	1214	100.0	678	12	AAAG95912
16	1214	100.0	678	3	AAAG95911
17	1214	100.0	678	3	AAAG95910
18	1214	100.0	678	7	AAAG95909
19	1214	100.0	678	8	AAAG95908
20	1214	100.0	678	8	AAAG95907
21	1214	100.0	678	14	AAAG95906
22	1214	100.0	678	9	AAAG95905
23	1214	100.0	678	10	AAAG95904
24	1214	100.0	678	11	AAAG95903
25	1214	100.0	678	11	AAAG95902
26	1214	100.0	678	8	AAAG95901
27	1214	100.0	678	10	AAAG95900
28	1214	100.0	678	13	AAAG95899
29	1214	100.0	678	13	AAAG95898
30	1214	100.0	678	6	AAAG95897
31	1214	100.0	678	6	AAAG95896
32	1214	100.0	678	12	AAAG95895
33	1214	100.0	678	4	AAAG95894
34	1214	100.0	678	4	AAAG95893
35	1214	100.0	678	4	AAAG95892
36	1214	100.0	678	6	AAAG95891
37	1214	100.0	678	12	AAAG95890
38	1214	100.0	678	12	AAAG95889
39	1214	100.0	678	12	AAAG95888
40	1214	100.0	678	6	AAAG95887
41	1214	100.0	678	6	AAAG95886
42	1214	100.0	678	6	AAAG95885
43	1214	100.0	678	4	AAAG95884
44	1214	100.0	678	13	AAAG95883
45	1214	100.0	678	13	AAAG95882
46	1214	100.0	678	6	AAAG95881
47	1214	100.0	678	11	AAAG95880
48	1214	100.0	678	11	AAAG95879
49	1214	100.0	678	14	AAAG95878
50	1214	100.0	678	14	AAAG95877
51	1214	100.0	678	6	AAAG95876
52	1214	100.0	678	6	AAAG95875
53	1214	100.0	678	14	AAAG95874
54	1214	100.0	678	14	AAAG95873
55	1214	100.0	678	14	AAAG95872
56	1214	100.0	678	14	AAAG95871
57	1214	100.0	678	13	AAAG95870
58	1214	100.0	678	12	AAAG95869
59	1214	100.0	678	13	AAAG95868
60	1214	100.0	678	13	AAAG95867
61	1214	100.0	678	12	AAAG95866
62	1214	100.0	678	11	AAAG95865
63	1214	100.0	678	11	AAAG95864
64	1214	100.0	678	3	AAAG95863
65	1214	100.0	678	10	AAAG95862
66	1214	100.0	678	10	AAAG95861
67	1214	100.0	678	11	AAAG95860
68	1214	100.0	678	11	AAAG95859
69	1214	100.0	678	12	AAAG95858
70	1214	100.0	678	12	AAAG95857
71	1214	100.0	678	10	AAAG95856
72	1214	100.0	678	3	AAAG95855
73	1214	100.0	678	6	AAAG95854
74	1214	100.0	678	3	AAAG95853
75	1214	100.0	678	3	AAAG95852
76	1214	100.0	678	3	AAAG95851
77	1214	100.0	678	4	AAAG95850



XX Stack JH, Whitney M, Cubitt AB, Pollak BA;  
PI MPI; 2001-468890/53.  
XX Destabilizing proteins in living cells, by coupling a target protein to  
XX PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin  
XX PT protein endoproteases, comprising two copies of the domain.  
XX  
XX Disclosure; Page 110; 171pp; English.  
XX  
XX The present invention relates to a method for destabilizing a target  
XX CC protein in a cell. The method comprises a linker moiety which operatively  
XX CC couples a target protein (a reporter moiety) to a linear multimerized  
XX CC destabilizing domain, which is non-cleavable by a -NH-ubiquitin protein  
XX CC endoproteases. The method is useful for detecting an activity such as  
XX CC protease, protein kinase or phosphoprotein phosphatase activity and is  
XX CC also useful for identifying novel assays for a wide range of post-  
XX CC translational activities, such as proteolysis, phosphorylation,  
XX CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,  
XX CC disulfide bond formation and ADP-ribosylation within cells. The  
XX CC recombinant DNA molecule of the invention is useful for creating  
XX CC transgenic animals useful as disease models and transgenic plants with  
XX CC improved disease resistance or other favourable traits. The present  
XX CC sequence is Discosoma sp. "red" anthozoa fluorescent protein, drFP483  
XX CC cDNA which is a natural fluorescent protein used as a reporter moiety in  
XX CC the exemplification of the invention

XX Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,3e-140 Length: 678  
Score: 1224.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-006-922a-12 (1-225) x AAD13053 (1-678)

QY 1 MetArgSerSerIysAlaValIleIleValGluPheMetArgPheIleValArgMetGluGly 20  
DB 1 ATGAGGCTCTTCCAAGATGTTATCAAGAGTTCATGAGTTAAAGTTCCCATGAAAGGA 60  
QY 21 ThrValAsnGlyIleGluPheGluIleGluGlyValGluGlyArgProGlyArgGly 40  
DB 61 ACCGTCATATGGGACGAGTTGAAATGAAAGGCAAGAGGAGGAGCCATACGAAGGC 120  
QY 41 HisAsnThrValIleValLeuValValThrIleGlyGlyProLeuProPheAlaTPAaPile 60  
DB 121 CACAAATACCGTAAAGCTTAAGTBAACCAAGGGGAGCCTTGGCATTTGGTGGAAATAT 180  
QY 61 LeuSerProGlnPheGlnIleGlySerIleValIleValIleValIleValIleValIle 80  
DB 181 TTGTCAACCAATTTGATGATGAAAGCAAGTATATCTCAAGACCCCTGCCGACATACCA 240  
QY 81 AspTyrIleValLeuSerPheProGluGlyPheIleThrGluArgValMetAsnPhGlu 100  
DB 241 GACTAATAAAAAGCTGATCTTCCGAAAGAAATTAAATGGAAGGGTCATGAACCTTTGAA 300  
QY 101 AspGlyValIleValIleThrValIleThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGCGCTCGTTACTGTACCAAGATTCAGCTTTGACAGATGCTGTTTCATCTAC 360  
QY 121 LysValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnIleValThr 140  
DB 361 AAGGTCAAGTTCAATGGCGTGAACCTTCTCCATGAGACCTGTTATGCAAAAGAAACA 420  
QY 141 MetGlyThrGluIleAsnThrGluArgLeuTyrProArgAspGlyValIleValGlyGlu 160  
DB 421 ATGGGCTGGGAAGCAGACCTGAGCGTTTGATCTCGTGAATGGCGTGTGAAGAGAG 480

QY 161 IleHisIleValIleValLeuValLeuValAspGlyIleGlyIleValIleValIleValIle 180  
DB 481 ATTCAATAAAGCTCGTAAGCTGAAGACGAGTGTCACTACCTAGTTCAATTCAAAAGTATT 540  
QY 181 TyrMetAlaIleValProValGlnLeuProGlyIleTyrIleValIleValAspSerIleAsp 200  
DB 541 TACATGGCAAGAACCTGTGCAAGTACCAAGGTCATCTAATGTTGACTCCAAATCGAAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValIleGluGlnIleGluArgThrGluGlyArg 220  
DB 601 ATTAACAAGCCACAAGACCAATCAATCTTGAAGCATGATGAAGAAGAACGAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCAATCTGTTCTCT 675  
RESULT 3  
ID AAD11142 standard; DNA; 678 BP.  
AC AAD11142;  
XX 24-SEP-2001 (first entry)  
XX  
XX Discosoma species Anthozoa fluorescent protein, drFP583 encoding DNA.  
XX  
XX Transmembrane potential; biological membrane; fluorescent ion; detection;  
XX test chemical screening; Anthozoa fluorescent protein; FP;  
XX transgenic organism; drFP583 protein; ds.  
XX  
XX Discosoma sp.  
XX  
XX WO200142211-A2.  
XX  
XX 14-JUN-2001.  
XX  
XX 12-DEC-2000; 2000WO-US033739.  
XX  
XX 13-DEC-1999; 99US-00459956.  
XX  
XX (REGC ) UNIV CALIFORNIA.  
XX  
XX Tsien RY, Gonzalez JB;  
XX MPI; 2001-457276/49.  
XX  
XX Determining electrical potential across a membrane in biological systems,  
XX PT comprises introducing two reagents, exposing the membrane to light and  
XX PT measuring the energy transfer.  
XX  
XX Disclosure; Page 150; 154pp; English.  
XX  
XX The patent discloses optical methods and compositions for determining  
XX CC transmembrane potential across biological membranes in living cells. The  
XX CC method of determining the electrical potential across a membrane in a  
XX CC biological system comprises introduction of two reagents, a first reagent  
XX CC comprising a hydrophobic fluorescent ion capable of redistributing from a  
XX CC first face of the membrane to a second face of the membrane in response  
XX CC to membrane potential change and a second reagent that label the first  
XX CC face or the second face of the membrane, which comprises a chromophore  
XX CC capable of undergoing energy transfer by either donating or accepting  
XX CC excited state energy to the fluorescent ion. The membrane is then exposed  
XX CC to excitation light and the energy transfer between the reagents is  
XX CC measured and related to the membrane potential. The method is useful for  
XX CC detecting changes in membrane potential in subcellular organelle  
XX CC membranes in biological systems. The method is used for screening of test  
XX CC chemicals for activity to modulate the activity of target ion channel.  
XX CC The invention also provides a transgenic organism comprising a first  
XX CC reagent that comprises a charged hydrophobic fluorescent molecule and a  
XX CC second reagent comprising a bioluminescent or naturally fluorescent  
XX CC protein. The present sequence is Discosoma species (red) DNA encoding an  
XX Anthozoa fluorescent protein (FP), drFP583

Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores: 2.3e-140 Length: 678  
 Pred. No.: 1214.00 Matches: 225  
 Score: 100.00% Conservat: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% FT  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-006-922A-12 (1-225) x AAD11142 (1-678)

```

Oy 1 MetATGSeSerLyAsnValIleLySGluPhMeTArGPhelyValArgMeGluGly 20
Db 1 ATGAGTCTTCCAAATACTTATCAAGAGTTCAAGAGTTTAAGGTTTCGCAAGAGA 60
Oy 21 ThrValaAngLyHISGluPhGluIleGluGlyGluGlyValArgProTyArgGly 40
Db 61 ACGGTCAATGGGCAAGAGTTTGAATAGAAAGCGAAGGAGGAGGCGCATACGAAGGC 120
Oy 41 HisAenThrValLyLeuLyValThrLySGlyGlyProLeuProPhaIaATPaspIle 60
Db 121 CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGAGCCCTTGCCATTGGCTGGATAT 180
Oy 61 LeuSerProGlnPhGlnTyGlySerLyValTyValLyHisProAlaAspIlePro 80
Db 181 TTGTACCAACCAATTCAGATAGGAAGCAAGATATATGCAACACCTCCGACATACCA 240
Oy 81 AspTyLyLyLeuSerPheProGluGlyPheLySTPGluArgValMetAsnPhGlu 100
Db 241 GACTATTAAGAGCTGTCTATCTCTGAAGGATTTAATGGGAAAGGTCATGAACCTTTGAA 300
Oy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyPheIleTy 120
Db 301 GACGGTGGCGTGTACTGTAAACCAAGATTCACAGATTGCAAGATGGCTGTTCACTAC 360
Oy 121 LySValLyPheIleGlyValAsnPhPheProSerAspGlyProValMetGlnLySlyThr 140
Db 361 AAGGTCAAGTCAATGGCTGTAACCTTCTCCGATGACCTGTATGCAAAAGAGACA 420
Oy 141 MetGlyTPGluAlaSerThrGluArgLeuTyProArgAspGlyValLeuLySGlyGlu 160
Db 421 ATGGCTGGGAAGCAGACATGAGCGTTGTATCTCCGTAAGCGGTGTAAGAGAGAG 480
Oy 161 IleHisLyValLeuLyLeuLyAspGlyGlyHisTyArgLeuValGluPhelySerIle 180
Db 481 ATTCTAAAGGCTCTGAAGCTGAAAGACGGTGGTCACTACCTAAGTAATCAAAAGTAT 540
Oy 181 TyrMeAlaLyLyLeuProValGlnLeuProGlyTyTyTyTyValAspSerLyLeuAsp 200
Db 541 TACATGGCAAGAAACCTGTGTGCACTACAGGGTACTACTATGTTGACTCCAAACTGAT 600
Oy 201 IleThrSerHisAenGluAspTyThrIleValGluGlnTyArgThrGluGlyArg 220
Db 601 ATTAACAACCCACAAGAGACTATCAATCGTTGAGCAGATGAAGAAAGACCGAGGAGCGC 660
Oy 221 HisHisLeuPheLeu 225
Db 661 CACCATCTGTCTCTT 675

```

RESULT 4

ABA95905 standard; cDNA; 678 BP.

ABA95905;

29-MAY-2002 (first entry)

Yeast optimised RFP encoding cDNA SEQ ID NO 1.

Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
 Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
 Escherichia coli; green fluorescent protein; biotechnology; gene; ss.

```

XX XX Anthozoa.
OS Key Location/Qualifiers
XX FT 1..678
XX CDS /tag= a
XX FT /product= "Yeast Optimised Red Fluorescent Protein"
XX DE20001395-U1.
XX PD 15-MAR-2001.
XX XX 27-JAN-2000; 2000DE-02001395.
XX XX 27-JAN-2000; 2000DE-02001395.
XX PR (GPCB-) GPC BIOTECH AG.
XX PA WPI; 2002-228394/29.
XX DR P-PSDB; ABB08834.
XX PT New DNA encoding red fluorescent protein, useful as marker in
XX PT biotechnology, has sequence optimized for expression in eukaryotes,
XX PT especially yeast or plants.
XX PS Claim 1; Fig 1; 19pp; German.
XX CC The invention relates to DNA (I) containing either sequence ABA95905 or
XX CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein
XX CC (YRFP). (I) are used to express red fluorescent protein (RFP) in
XX CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and
XX CC plants, especially dicotyledonous plants including Nicotiana tabacum or
XX CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,
XX CC especially Escherichia coli. RFP is useful in the same way as green
XX CC fluorescent protein but is more generally applicable in modern
XX CC biotechnology. (I) are optimised for expression in yeast and so generate
XX CC RFP at higher levels with stronger fluorescence and thus lowers the
XX CC detection limit and gives a better signal-to-noise ratio
XX SQ Sequence 678 BP; 198 A; 147 C; 159 G; 174 T; 0 U; 0 Other;

Alignment Scores: 2.3e-140 Length: 678
Pred. No.: 1214.00 Matches: 225
Score: 100.00% Conservat: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% FT
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x ABA95905 (1-678)
Oy 1 MetATGSeSerLyAsnValIleLySGluPhMeTArGPhelyValArgMeGluGly 20
Db 1 ATGAGTCTTCCAAATACTTATCAAGAGTTCAAGAGTTTAAGGTTTCGCAAGAGA 60
Oy 21 ThrValaAngLyHISGluPhGluIleGluGlyGluGlyValArgProTyArgGly 40
Db 61 ACGGTCAATGGGCAAGAGTTTGAATAGAAAGCGAAGGAGGAGGCGCATACGAAGGC 120
Oy 41 HisAenThrValLyLeuLyValThrLySGlyGlyProLeuProPhaIaATPaspIle 60
Db 121 CACAACTGTCAAGAGTTTGAAGGTGATCTAAGGTTGATGATGATGATGATGATGAT 180
Oy 61 LeuSerProGlnPhGlnTyGlySerLyValTyValLyHisProAlaAspIlePro 80
Db 181 TTGTCTCCAAATTCAGATAGGAAGCAAGATATATGCAACACCTCCGACATACCA 240
Oy 81 AspTyLyLyLeuSerPheProGluGlyPheLySTPGluArgValMetAsnPhGlu 100
Db 241 GACTACAGAGAGTGTCTCTTCCGAGAGTTTCAAGTGGGAAAGGATCATCAACTTCCA 300
Oy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyPheIleTy 120

```

Db 301 GACGGTGTGTTGTTACTGTTACTCAAGACTCCTCTCTGCAAGACGGTTGTTTCATCTAC 360  
 Qy 121 LysValIysPheIleGIYValAsnPhenProSerAspGIYProValMetGlnYsIysThr 140  
 Db 361 AAGGTCAATTCATCGGTCTCACTCCCATCTGAACGGTCCAGTATGCAAAAGAAAGACT 420  
 Qy 141 MetGIYTrpGlnAlaSerThrGluArgLeuTYrProArgAspGIYValIleuysGIYgln 160  
 Db 421 ATGGGTGGGAAGCTTCTACCAACGTTGTGACCCCAAGACGGTGTCTTGAAAGGGTGA 480  
 Qy 161 ILeHleValAlaLeuYsIleuYsAspGIYgIYHleTYrLeuValGlnPhelYsSerIle 180  
 Db 481 ATCCACAAAGCCCTTGAAAGTGAAGACGGGTGGTCACTACTGTCATTCAGATTATAC 540  
 Qy 181 TYrMetAlaIysLeuProValGlnLeuProGIYTYrTYrTYrValAspSerIysLeuAsp 200  
 Db 541 TACATGGCTTAAGAGCACTCCCAATTGCCAGTTTACTACTAGCTTGACTTAAGTTGAC 600  
 Qy 201 ILeThrSerHlaAsnGlnuAspTYrThrIleValGlnGlnTYrGlnuArgThrGlnuArg 220  
 Db 601 ATCACCTCTCACAAACGAAGCTACACTATCGTCGAACAATACGAACGTACTGAAGGTAGA 660  
 Qy 221 HleHleLeuPhelu 225  
 Db 661 CACCACTGTGTTCTTG 675  
 RESULT 5  
 ABA95922  
 ID ABA95922 standard; DNA; 678 BP.  
 AC ABA95922;  
 XX  
 DT 29-MAY-2002 (first entry)  
 XX  
 Yeast optimised RFP related DNA SEQ ID NO 19.  
 DE Yeast optimised RFP related DNA SEQ ID NO 19.  
 XX  
 Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
 KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
 KM Escherichia coli; green fluorescent protein; biotechnology; gene; ss.  
 XX  
 Anthozoa.  
 OS  
 Key Location/Qualifiers  
 CDS 1..678  
 FT /tag= a  
 FT /product= "Yeast optimised Red Fluorescent Protein"  
 FT  
 DE20001395-U1.  
 XX  
 PN  
 PD 15-MAR-2001.  
 PF 27-JAN-2000; 2000DE-02001395.  
 XX  
 PR 27-JAN-2000; 2000DE-02001395.  
 XX  
 (GPCR-) GPC BIOTECH AG.  
 XX  
 MPI; 2002-228394/29.  
 DR  
 PT New DNA encoding red fluorescent protein, useful as marker in  
 PT biotechnology, has sequence optimized for expression in eukaryotes,  
 PT especially yeast or plants.  
 XX  
 PS Disclosure, Page 14-15; 19pp; German.  
 CC The invention relates to DNA (I) containing either sequence ABA95905 or  
 CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
 CC (YRFP). (I) are used to express red fluorescent protein (RFP) in  
 CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
 CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
 CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
 CC especially Escherichia coli. RFP is useful in the same way as green  
 CC fluorescent protein but is more generally applicable in modern

CC biotechnology. (I) are optimised for expression in yeast and so generate  
 CC RFP at higher levels with stronger fluorescence and thus lowers the  
 CC detection limit and gives a better signal-to-noise ratio. The present  
 CC sequence is that of a polynucleotide encoding the yeast optimised RFP,  
 CC useful to the invention  
 XX

SQ Sequence 678 BP; 202 A; 118 C; 159 G; 199 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	2,3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x ABA95922 (1-678)

Qy 1 MetArgSerSerIysAsnValIleYsGlnPhenMetArgPheIysValArgMetGlnuY 20  
 Db 1 ATGAGATCTTCTTAAGAACGTTATTAAGAAATTCATGAGATTCAGGTTGAATGGAAGT 60  
 Qy 21 ThrValAsnGIYHleGlnPhelGlnIleGlnuYgIYgIYgIYArgProTYrGlnuY 40  
 Db 61 ACTGTTAACGGTCAAGAAATTCGAATTCGAAGGTGAAGGTGAAGGTATGACCATACGAAGT 120  
 Qy 41 HleAsnThrValIysLeuYsValThrIysGIYgIYProLeuProPheHleATrpaPheIle 60  
 Db 121 CACAACACTGTTAAGTGAAGGTTCATTAAGGGTGGTCCATTCGCAATTCCTGGGACATT 180  
 Qy 61 LeuSerProGlnPhelGlnTYrGlnYsSerIysValTYrValIysHleProAlaAspIlePro 80  
 Db 181 TTGTCTCCCAATCCAAATCGATTCATTAAGTTTACGTTTAAGCAACCACTGACATTCGA 240  
 Qy 81 AspTYrIysLeuSerPheProGlnuYpHeIYsTrpGlnuArgValMetAsnPhelGln 100  
 Db 241 GACTCAAGAAAGTGTCTTCCAGAAAGTTCAAGTGGAAAGATTATGAACCTTCGA 300  
 Qy 101 AspGIYgIYValIysValThrValThrGlnAspSerSerIleGlnAspGIYCySpHeIleTYr 120  
 Db 301 GACGGTGTGTTGTTACTGTTACTCAAGACTCTCTCTTTCGAAGACGGTGTTCATTTAC 360  
 Qy 121 LysValIysPheIleGIYValAsnPhenProSerAspGIYProValMetGlnYsIysThr 140  
 Db 361 AAGGTCAATTCATCGGTCTCACTCCCATCTGAACGGTCCAGTATGCAAAAGAAAGACT 420  
 Qy 141 MetGIYTrpGlnAlaSerThrGluArgLeuTYrProArgAspGIYValIleuysGIYgln 160  
 Db 421 ATGGGTGGGAAGCTTCTACCAACGTTGTGACCCCAAGACGGTGTCTTGAAAGGGTGA 480  
 Qy 161 ILeHleValAlaLeuYsIleuYsAspGIYgIYHleTYrLeuValGlnPhelYsSerIle 180  
 Db 481 ATCCACAAAGCCCTTGAAAGTGAAGACGGGTGGTCACTACTGTCATTCAGATTATAC 540  
 Qy 181 TYrMetAlaIysLeuProValGlnLeuProGIYTYrTYrTYrValAspSerIysLeuAsp 200  
 Db 541 TACATGGCTTAAGAGCACTCCCAATTGCCAGTTTACTACTAGCTTGACTTAAGTTGAC 600  
 Qy 201 ILeThrSerHlaAsnGlnuAspTYrThrIleValGlnGlnTYrGlnuArgThrGlnuArg 220  
 Db 601 ATTACTTCTCACAAACGAAGCTACACTATTTGGAACAATACGAAGAACTGAAGGTAGA 660  
 Qy 221 HleHleLeuPhelu 225  
 Db 661 CACCACTGTGTTCTTG 675  
 RESULT 6  
 ABA95921  
 ID ABA95921 standard; DNA; 678 BP.  
 AC ABA95921;  
 XX  
 DT 29-MAY-2002 (first entry)

[illegible]

Dd	18	TTTGACCAACAATTTCAGATATGAAAGCAAGGATATATGCAAGCAACCTGCAGATACCA	240
Qy	81	ASPTTYLysLysLeuSerPheProGluGlyPheLYSTProGluValGluMetAsnPheGlu	1000
Dd	241	GACTATATAAAACCTGTCAATTTCTCGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA	3000
Qy	101	ASPGLYGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	1200
Dd	301	GACGGTGGGGTGGTACTACTGTAAACCAAGATTCGAGTTTGACGAGATGGCTGTTTCATCTAC	3600
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	1400
Dd	361	AAGGTCAGAGTTCATTTGGGGTGAACTTTCTCCATGGAAGCCTGTTTATGCAAAAGAAACA	4200
Qy	141	MetGlyTrpGlnAlaSerThrGlnLysLeuTyrProArgAspGlyValLleuLysGlyL	1600
Dd	421	ATGGCTGGGAAAGCCAGCACTGAGGGCTTTTGATCTCTGTATGGCCGTGTGAAAGGAGAG	4800
Qy	161	IleHisLysValLleuLysLeuLysAspLYGlyHisLysTyrLeuValGluPheLysSerIle	1800
Dd	481	ATTCATTAAGGCTCTGAGGCTGAAGAAGCGGTGATTAACCTAGTTGAATTCAAAAGTAATT	5400
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	2000
Dd	541	TACATGGCAAAAGAACCTGTGACGTACCAAGGTACTATATGTGACTCCAAACTGGAT	6000
Qy	201	IleThrSerHisAspGlyAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	2200
Dd	601	ATPACAGGCAACAAGAAAGACTATACATCGTTGAGCGATGTAAAGAAACCGAGGAGCGC	6600
Qy	221	HisHisLeuPheLeu 225	
Dd	661	CACCAATCTGTTCCTT 675	
RESULT 7			
ABA95920			
ID	ABA95920 standard; RNA; 678 BP.		
XX	ABA95920;		
AC	29-MAY-2002 (first entry)		
DT			
XX			
DE	Yeast optimised RFP encoding RNA SEQ ID NO 16.		
KW	Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;		
KW	Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;		
KW	Escherichia coli; green fluorescent protein; biotechnology; ss.		
OS	Anthozoa.		
XX			
XX	DE20001395-U1.		
XX	15-MAR-2001.		
PD			
XX	27-JUN-2000; 2000DE-02001395.		
XX	27-JUN-2000; 2000DE-02001395.		
PR	(GPCB-) GPC BIOTECH AG.		
XX			
PA	WPI; 2002-228394/29.		
DR			
XX			
XX			
PT	New DNA encoding red fluorescent protein, useful as marker in		
PT	biotechnology, has sequence optimized for expression in eukaryotes,		
PT	especially yeast or plants.		
XX			
XX	Disclosure; Page 13; 19pp; German.		
CC	The invention relates to DNA (I) containing either sequence ABA95905 or		
CC	sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein		
CC	(YRFP). (I) are used to express red fluorescent protein (RFP) in		
CC	eukaryotes, especially yeast, especially Saccharomyces cerevisiae and		
CC	plants, especially dicotyledonous plants including Nicotiana tabacum or		



CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
 CC especially Escherichia coli. RFP is useful in the same way as green  
 CC fluorescent protein but is more generally applicable in modern  
 CC biotechnology. (1) are optimised for expression in yeast and so generate  
 CC RFP at higher levels with stronger fluorescence and thus lowers the  
 CC detection limit and gives a better signal-to-noise ratio. The present  
 CC sequence is that of an RNA sequence corresponding to the yeast optimised  
 CC RFP encoding cDNA given as SEQ ID NO 1 (ABA95905)

XX  
 SQ Sequence 678 BP; 198 A; 147 C; 159 G; 0 T; 174 U; 0 Other;

#### Alignment Scores:

Pred. No.:	2,3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x ABA95920 (1-678)

```

QY      1 MetArGSeSerLyAsnValIleLySGluPhMetArGPhelyValArGmetGluGly 20
DB      1 AUGAGAUCCUCCAAAGACGUCACAAAGAAUUCAGAUUCAGAGUUAAGAGAGG 60
QY      21 ThrValaAngLyHtAgIupheGluIleGluGlyGluGlyGluGlyGluGlyGluGly 40
DB      61 ACUGUUAACGUGUACGAAUUCGAAUUCGAAAGGUGAAAGUAGACCAUUCGAAAGGU 120
QY      41 HtAenThrValLyLeuLyValThrLyGlyGlyProLeuProPhelaTTPaPille 60
DB      121 CACAAACAGUCGAAAGUUGAAGUACUAAAGGUGUCUACUUCGUAUUCGUGAGACUC 180
QY      61 LeuSeProGluPhheGluInTyGlySerLyValTyValLyHtAspProAlaPillePro 80
DB      181 UUUUCCCAUAUUCGAGUUCUAGGUCUAGGUCUAGGUCUAGGUCUAGGUCUAGGUCU 240
QY      81 AsPtyrLyLyLeuSeRPhheProGluGlyPhelyTTPGluArgValMetAenPhelGlu 100
DB      241 GACUACAAAGAAUGUUCUCCCAAGAGUUCUAAAGGUGAAAGUACUAGAACTUCCAA 300
QY      101 AsPGLyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCySPhelleTyr 120
DB      301 GACGUGUGUUGUUCUAGUACUACGACUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCU 360
QY      121 LyValLyAspHeTleGlyValAenPhheProSerAspGlyProValMetGluLySlyThr 140
DB      361 AAGGUCAAAGUUCAGGUGUCACUUCUCCAGGUCUCCAGGUCUCCAGGUCUCCAGGUCU 420
QY      141 MetGlyTTPGluAspSerThrGluArgLeuTyrProArgAspGlyValLeuLyGlyGlu 160
DB      421 AUGGUGUGGAGAGUUCUACGAAAGUUCUCCCAAGAGCGGUGUCCUUGAAGGUGAA 480
QY      161 HtAenLyValLyLeuLyLeuLyAspGlyGlyHtSlyrLeuValGluPhelySerIle 180
DB      481 AUCCACAAAGCCUUGAAGUAGAGAGGUGUCUACUACUUGGCGCAUUCAGAGUCCAUUC 540
QY      181 TyrMetCtAlaLySlySProValGluLeuProGlyTyrTyrTyrValAaSPserLyLeuAsp 200
DB      541 UACUGUGCUAAGAGCCAGUCCAUUCGAGUUAACAUAUGUGUACUCCUAGUAGUAGAC 600
QY      201 HtAenThrValaAngLyAspLyTThrIleValGluGluInTyGluArgThrGluGly 220
DB      601 AUCACCCUCCAAAGAGAAACUACAUUCUGAACAUAUCGAAAGUACUAGAGUAGA 660
QY      221 HtAenLyLeuPhelu 225
DB      661 CACCACUUGUUCUG 675

```

RESULT 8  
 AAD46278  
 ID AAD46278 standard; DNA; 678 BP.  
 XX

```

AC      AAD46278;
XX      27-DEC-2002 (first entry)
XX      Discosoma sp. drFP583 (NFP-6) wild-type protein encoding DNA.
DE      Discosoma sp. drFP583 (NFP-6) wild-type protein encoding DNA.
XX      Fluorescent protein; chromoprotein; protease cleavage assay; filter;
KW      fluorescence activated cell sorting application; fluorescent timer;
KW      biosensor; fluorescence resonance energy transfer application; FRER;
KW      colouring agent; recombinant DNA application; analyte detection assay;
KW      sunscreen; second messenger detector; drFP583 protein; NFP-6; gene; ds.
XX      Discosoma sp.
OS      Discosoma sp.
XX      Key Location/Qualifiers
XX      CDS 1..678
XX      FT /*tag= a
XX      FT /product= "drFP583 wild-type protein"
XX      PN WO200268459-A2.
XX      PD 06-SEP-2002.
XX      PF 20-FEB-2002; 2002WO-US005749.
XX      PR 21-FEB-2001; 2001US-0270983P.
XX      PR 04-DEC-2001; 2001US-00006922.
XX      PA (CLON-) CLONTECH LAB INC.
XX      PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;
XX      DR MPI; 2002-591654/74.
XX      DR P-PSDB; AAE28833.
XX      PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant
XX      PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for
XX      PT analyte detection assays or fluorescence activated cell sorting
XX      PT applications.
XX      PS Disclosure; Page 70; 80pp; English.
XX      SS
XX      CC The invention relates to nucleic acid molecules encoding non-aggregating
XX      CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are
XX      CC useful in analyte detection assays, as colouring agents, as markers in
XX      CC recombinant DNA applications, as sunscreens or filters, in fluorescence
XX      CC resonance energy transfer (FRER) applications, as biosensors in
XX      CC prokaryotic and eukaryotic cells, in screening assays, as second
XX      CC messenger detectors, in fluorescence activated cell sorting applications,
XX      CC in protease cleavage assays or as fluorescent timers. The present
XX      CC sequence is a DNA encoding Discosoma sp. drFP583 (NFP-6) wild-type
XX      CC protein of the invention
XX      SQ Sequence 678 BP; 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;

```

#### Alignment Scores:

Pred. No.:	2,3e-140	Length:	678
Score:	1214.00 <td>Matches:</td> <td>225</td>	Matches:	225
Percent Similarity:	100.00% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	100.00% <td>Mismatches:</td> <td>0</td>	Mismatches:	0
Query Match:	100.00% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AAD46278 (1-678)

```

QY      1 MetArGSeSerLyAsnValIleLySGluPhMetArGPhelyValArGmetGluGly 20
DB      1 AUGGCTCTCCAAAGACGTCATCAAGAGATTGATGGCTTCAAGTGGCGCATGAGAGGC 60
QY      21 ThrValaAngLyHtAgIupheGluIleGluGlyGluGlyGluGlyGluGlyGluGly 40
DB      61 ACCGTGAACGCGCAAGATTCAAGATCAGGAGGCGAGGCGCGCCCTCAAGAGAGGC 120

```

```

QY 41 HisAenThrValIyLeuIySerThrIySGIyGIyProLeuProPheAlaTrpAspIle 60
DB 121 CACAACACCGCTGAAGCTGAAGTGAACCAAGGGGGCCCTTCCTTCGGGACATC 180
QY 61 LeuSerProGlnPheGlnTyrgIySerIyValIyTrValIySHIaProAlaAspIlePro 80
DB 181 CTGTCCCCCCCACTTCCAGTACCGGCTCCCAAGGTGTAAGTAAACACCCGCCACATCCCC 240
QY 81 AspTyIyIyLeuIyLeuSerPheProGlnIyPheIySTrPGIyAryValMetAsnPheGlu 100
DB 241 GACTACCAAGAACTGTCTTCCCGAGGGCTTCAGTGGAGCGGTGATGAACCTTCGAG 300
QY 101 AaPGIyGIyValIyValIyThrValIyThrgIyAaPSeSerIeGlnIaAspGIyCyapPheIleTy 120
DB 301 GACGGCGGGGTGTGACCGCTGACCCAGACCTCTCCCTCCAGAGACGGCTGCTTCACTTAC 360
QY 121 IySValIyPheIleGIyValIyAaPheProSerAaPGIyProValMetGlnIySlyThr 140
DB 361 AAGTGAAGTTCATCGGCGCTGAACCTTCCCTCCGACGGCCCGTGTATGCAAGAAAGACC 420
QY 141 MetGIyTrPGIyAaIaSerThrGlnIyrgIeUyTrProAryAaAspGIyValIleuIySGIyGlu 160
DB 421 ATGGGCTGTGGAGGCTTCCACCAAGGCGCTGTACCCCGGACGGCGTGTGAAGGGCGAG 480
QY 161 ILeHISlyAAlaLeuIyLeuIySaaPGIyGIyHISlyTrLeuValIgluPheIySerIle 180
DB 481 ATCCCAAGAGCCCTGAAGCTGAAGACGGCGGCCACTACTGTGTGAAGTTCAGATCCATC 540
QY 181 TyrMetAlaIyIyPheProValIgluIeupProGIyTyTrTyTrValIaPSeSerIySleuAaP 200
DB 541 TACATGTGCGCAAGAGCCGTGACGCTGCCCGGCTACTACTACGTGTGATCCCAAGCTGGAC 600
QY 201 ILeThrSerHISaSGIyAaPTrThrIleValIgluIyGlnIyThrgIyAaPheIyGluIyAaG 220
DB 601 ATCACCCTCCCAACAGAGGACTACACCATGTGTGAGCACTACAGGACCCAGAGGGCGCG 660
QY 221 HISHISLeuPheIeU 225
DB 661 CACCACTCTTCTCTG 675

```

RESULT 9  
AAD28207  
ID AAD28207 strand; DNA; 678 BP.  
XX  
AC AAD28207;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Diacoboma sp. humanised wild-type Anthozoa protein drfp583 DNA.  
XX  
KM Fluorescent timer protein; protein movement; translocation; trafficking;  
KM promoter activity; gene expression; transgenic plant; gene modification;  
KM protein age; anthozoa protein; drfp583; ds.  
XX  
OS Diacoboma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /tag= a  
FT /product= "humanised wild-type Anthozoa protein drfp583"  
XX  
PN MO200196373-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019097.  
XX  
PR 14-JUN-2000; 2000US-0211607P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Pradkov AF, Teresikh A;  
XX

```

DR WPI; 2002-154595/20.
DR P-PSDB; AAB17540.
XX
PT New fluorescent timer proteins comprising an emission spectrum that
PT changes over time from a first wavelength to a second wavelength, useful
PT for monitoring intracellular protein movement, translocation, trafficking
PT or stability.
XX
PS Example 1; Fig 1; 89pp; English.
XX
CC The invention relates to a fluorescent timer protein having an emission
CC spectrum that changes over time after synthesis from a first wavelength
CC to a second wavelength. The fluorescent timer proteins are useful in
CC monitoring the activity of a promoter, determining the age of a protein,
CC identifying an agent that modulates the activity of a promoter and in
CC enriching a population of cells comprising a fluorescent timer protein.
CC The fluorescent timer proteins are also useful for assessing gene
CC expression during development of a multicellular organism or during
CC cellular differentiation, in response to a drug or other inducer of
CC promoter activity, as a reporter to serve as a read-out of promoter
CC activity, monitoring intracellular protein movement or translocation,
CC protein trafficking, or protein stability, to investigate temporal
CC aspects of the activity of a regulatory element, for determining cell
CC fate during development and organ remodeling, in spatial and temporal
CC visualisation of newly synthesised proteins and accumulated proteins, and
CC in distinguishing between newly formed and pre-existing structures, e.g.
CC membrane junctions and extracellular matrix components. The fluorescent
CC timer proteins may further be used to investigations where photobleaching
CC techniques are employed, as detectable labels, as selectable markers, as
CC bioensors in prokaryotic and eukaryotic cells, in protease cleavage
CC assays, and as second messenger detectors. The nucleic acids can be used
CC to generate transgenic, non-human plants or animals or site-specific gene
CC modifications in cell lines. The present sequence is a DNA encoding
CC Diacoboma sp. humanised wild-type Anthozoa protein drfp583 used for
CC generating fluorescent proteins
XX
SQ Sequence 678 BP; 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;

```

Alignment Scores:  
Pred. No.: 2,3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-006-922A-12 (1-225) x AAD28207 (1-678)

```

QY 1 MeArgSerSerIySaaNValIleIySGIyPheMetAryPheIyValIyrgMetGIyGlu 20
DB 1 ATGGCTCTCTCCCAAGAAAGTCAAGAGTTCATGCGCTTCAAGGTGCGGATGAGGGG 60
QY 21 ThrValAaNGIyHISGlnPheGlnIleGIyGluIyGluIyAaPProTyGIyGlu 40
DB 61 ACCGTGAACGGCCACAGATTCAGATCGAGGGCGAGGGCGGCCCTTCCACAGAGGAC 120
QY 41 HisAenThrValIyLeuIyValIyThrIySGIyGIyProLeuProPheAlaTrpAspIle 60
DB 121 CACAACACCGTGAAGCTGAAGTGAACCAAGGGGGCCCTTCCTTCGGGACATC 180
QY 61 LeuSerProGlnPheGlnTyrgIySerIyValIyTrValIySHIaProAlaAspIlePro 80
DB 181 CTGTCCCCCCCACTTCCAGTACCGGCTCCCAAGGTGTAAGTAAACACCCGCCACATCCCC 240
QY 81 AspTyIyIyLeuIyLeuSerPheProGlnIyPheIySTrPGIyAryValMetAsnPheGlu 100
DB 241 GACTACCAAGAACTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAACCTTCGAG 300
QY 101 AaPGIyGIyValIyValIyThrValIyThrgIyAaPSeSerIeGlnIaAspGIyCyapPheIleTy 120
DB 301 GACGGCGGGGTGTGACCGCTGACCCAGACCTCTCCCTCCAGAGACGGCTGCTTCACTTAC 360
QY 121 IySValIyPheIleGIyValIyAaPheProSerAaPGIyProValMetGlnIySlyThr 140

```

DB 361 AAGGTGAATTATCGCGCTGAACTTCCCTCCGACGGCCCCGTGATGCAAGAAAGACC 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuylsGlyIu 160  
DB 421 ATGGGCTGGAGGCTTCCACCGAGCCCTGTACCCCGGACGCGCTGTGAAAGGCGAG 480  
QY 161 ILehIleValAlaLeuylsLeuylsAspGlyGlyYhiIeTyrIleuValGluPheylsSerIle 180  
DB 481 ATTCACAAAGGCGCTGAAGCTGAAGGACGCGGCGCACTACTGCTGAGTTCAAGTCCATC 540  
QY 181 TyrMetAlaIleuylsPProValGlnLeuProGlyTyrTyrTyrValAspSerIleuAsp 200  
DB 541 TACATGGCCAAAGAACCCGTCGACGCTGCCCTACTACTGACTGACCTTCACAGCTGAGAC 600  
QY 201 IleThSerHisAnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAGACGACTACACCACTGTGAGACATGACAGCGCACCGAGGCGCG 660  
QY 221 HLehIleuPheleu 225  
DB 661 CACCACTGTCTCTG 675  
RESULT 10  
ADC24127  
ID ADC24127 standard; DNA; 678 BP.  
XX  
XX ADC24127,  
XX  
XX 18-DEC-2003 (first entry)  
DT  
DE Discosoma wild-type red fluorescent protein DNA #1.  
XX  
XX Discosoma red fluorescent protein; Dared; AB interface; AC interface;  
KM fluorescence protein variant; transcription induction detection;  
KM fluorescence energy resonance transfer; FRBT; protein kinase;  
KM protein phosphatase; ion indicator; ds.  
XX  
XX Discosoma.  
OS  
XX US2003058935-A1.  
PN  
XX 27-MAR-2003.  
PD  
XX 10-APR-2002; 2002US-00121258.  
PF  
XX 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.  
XX  
XX (TSIE/) TSIE R Y.  
PA (CAMP/) CAMPBELL R E.  
XX  
PI Tsien RY, Campbell RE;  
PI  
PI WPI; 2003-743764/70.  
DR P-PSDB; ADC24126.  
XX  
XX Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.  
XX  
XX Example 1; SEQ ID NO 2; 67bp; English.  
XX  
XX The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (Dared) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type Dared sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the Dared variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (i) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant

CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcription, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca<sup>2+</sup> Zn<sup>2+</sup>, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This sequence encodes Discosoma wild-  
CC type red fluorescent protein.  
XX  
XX SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;  
XX  
XX Alignment Scores:  
Pred. No.: 2,3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0  
US-10-006-922a-12 (1-225) x ADC24127 (1-678)  
QY 1 MetArgSerSerIleValAnValIleValGluPheMetArgPheIleValArgMetGluGly 20  
DB 1 ATGAGGCTCTCCCAAGATGTTATCAAGAGTTGATGAGGTTTAAAGTTGCGCATGAGAAAGA 60  
QY 21 ThrValAsnGlyYhiIeGluPheGluIleGluGluGluGluGluGluGluGluGluGluGlu 40  
DB 61 ACGGTCAATGGGACGACGATTTGAAATGAAAGGCGAAAGGAGGAGGAGGCCATACGAGAGC 120  
QY 41 HLehAnThrValIleuylsLeuylsValIleuylsGlyIleuylsProLeuPheAlaTyrAspIle 60  
DB 121 CACAAATACCGTAAGCTTAAGTTAACCAAGGGGGACCTTTGCAATTTGCTGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerIleValIleuylsIleuylsProIleAspIlePro 80  
DB 181 TTGTACACCAATTTTCAGTATGAGAACCAAGTATATGTCAAGCACCTGCGCATACCA 240  
QY 81 AspTyrIleuylsLeuSerPhePProGluGlyPheIleTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTATATAAAGCTGTCATTTCTGAAAGATTTAAATGGGAAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValIleuylsValIleuylsThrValIleuylsAspSerIleuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCGTACTGTAACCCAGATTTCCAGTTGCGAGATGGCTTTTCATCTAC 360  
QY 121 ILeuValIleuylsPheIleGlyValIleuylsAsnPheProSerAspGlyProValMetGluIleuylsThr 140  
DB 361 AAGGTCAAGTTCAATTGGCGTGAACCTTTCTCCGATGACCTGTATGCAAAAGAAAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuylsGlyIu 160  
DB 421 ATGGGCTGGAGGACCGACACTGAGCGCTTGTATCTCTGTATGCGCGTGGAAAGGAGAG 480  
QY 161 ILehIleValAlaLeuylsLeuylsAspGlyGlyYhiIeTyrIleuValGluPheylsSerIle 180  
DB 481 ATTCATAGGCTCTGAAGCTGAAGGACGTCGTCATTAAGTTGAATTCAAAGATATT 540  
QY 181 TyrMetAlaIleuylsPProValGlnLeuProGlyTyrTyrTyrValAspSerIleuAsp 200  
DB 541 TACATGGCAAGAGGCTGTGACGCTACCAAGGTAATGATGTGTGACTCCAAACTGGAT 600  
QY 201 IleThSerHisAnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220

```
DB      601 ATTAACAGCCACACAGACCTATCAATCGTTGACGATATGAAGAACCGAGGAGCGC 660
QY      221 HlshlsluPhelu 225
DB      661 CACCATCTGTCTCTT 675

RESULT 11
ADP70404 standard; cDNA; 678 BP.
AC      ADP70404;
XX
XX      12-FEB-2004 (first entry)
DE      Discosoma wild-type GFP variant cDNA SeqID27.
XX
XX      ligand; orphan receptor protein; fusion protein; fluorescent protein;
XX      cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
XX      GFPuv; Enhanced GFP; EGFP; gene; ss.
OS      Discosoma sp.
XX
XX      WO2003071272-A1.
XX      28-AUG-2003.
XX      21-FEB-2003; 2003WO-JP001901.
XX      22-FEB-2002; 2002JP-00045728.
XX      23-JUL-2002; 2002JP-00213949.
XX      11-OCT-2002; 2002JP-00298237.
XX      (TAKE ) TAKEDA CHEM IND LTD.
XX
XX      Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
XX      MPI: 2003-697654/66.
XX      P-PSDB; ADP70403.
XX
XX      Transformation of cells with a fusion protein of an orphan receptor
XX      protein with a fluorescent protein useful for identification of ligands
XX      to the orphan receptor.
XX
XX      Disclosure; SEQ ID NO 27; 594bp; Japanese.
XX
XX      This invention relates to a novel method of identifying ligands to an
XX      orphan receptor protein which comprises transforming cells with DNA
XX      encoding a fusion protein of the orphan receptor with a fluorescent
XX      protein, so that the fusion protein is expressed in the cells (or cell
XX      membranes isolated from them) and contacting the cells with the potential
XX      ligand to be tested. A suitable fluorescent protein for incorporation in
XX      the fusion protein is green fluorescent protein (GFP), for example GFP-1,
XX      wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
XX      identification of ligands binding to an orphan receptor protein.
XX
XX      Sequence 678 BP; 204 A; 129 C; 179 G; 166 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      2.3e-140      Length:      678
Score:      1214.00      Matches:      225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB:      10      Gaps: 0

US-10-006-922A-12 (1-225) x ADP70404 (1-678)
QY      1 MetAAGSerSerLyAaenVal11lelyGluPhuMetAaRphelyValAArgMetGluGly 20
DB      1 ATGAGGTCCTTCAGAGATGTATCAAGAGATTCAAGAGTTTAAGGTTGCAATGAGAGCA 60
QY      21 ThrValaenGlyHlshlsluPhelu1leGluGlyGluGlyAArgProLyArgGluGly 40
```

```
DB      61 ACCGTCATGGGCAAGCATTTTGAATAATAGAGCGAAGAGAGGGGAGCGCATTACGAAGCG 120
QY      41 HlsaenThrValLyLeuLyValThrLyGlyGlyProLeuProPhelATrpAplle 60
DB      121 CACATATACCGTAAGCTTAAGGTAAACCAAGGGGAGCCTTGGCATTGCTTGCGATATT 180
QY      61 LeuSerProGlnPhelGlnTrpGlySerLyValTyValLyHlsAProAlaAsp1lePro 80
DB      181 TTGTACCAACAATTCATGATGAGACCAAGATATATGTCACACCTGCGCACTTACCA 240
QY      81 AspTyLyLyLeuSerPheProGluGlyPheLySTrPGluArgValMetAaPhelu 100
DB      241 GACTATTAATAAGCTCTCATTTCTCTGAAGATTAAATGGAAAGGTCATGAACCTTTGA 300
QY      101 AaPGlyLyValValThrValThrGlnAaPseSerLyGlnAaPGlyCyAaPhe1leTy 120
DB      301 GACGCTGGCGCTGTTACTGTAACCAAGATTCCAGTTTGACAGGATGCGCTGTTCACTTAC 360
QY      121 LySValLyPheHlshlsluValAaPhProSerAaPGlyProValMetGlnLySlyThr 140
DB      361 AAGTCAGTTCATTGCGGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAAAGACA 420
QY      141 MetGlyTrPGluAaSerThrGluArgLeuTyTrProAaAaPGlyValLeuLyGlyGlu 160
DB      421 ATGGGCTGGGAAGCAAGCACTGAGCGCTTGTATCTCTGATGGCGTGTGAAAGAGAG 480
QY      161 HlshlsluValAaLeuLyLeuLyAaAaPGlyGlyHlsTyLeuValGluPhelyAaSer1le 180
DB      481 ATTCATTAAGGCTCTGAAAGCTGAAGACGCTGCTCATTTCAATTGAATTCAAAAGTATT 540
QY      181 TyrMetAlaLyLySProValGlnLeuProGlyTyTyTyTyValAaSPSerLySLeuAaP 200
DB      541 TACATGGCAAGAAAGCCTGTGACCTACAGAGGATCTATGTTGACTCCAAACTGAT 600
QY      201 IleThrSerHlsAaenGluAaSPTyTrThr1leValGluGlnTyrgLuArgThrGluGlyArg 220
DB      601 ATTAACAGCCACACAGAGACCTATCAATCGTTGACGATGAAGAACCGAGGAGCGC 660
QY      221 HlshlsluPhelu 225
DB      661 CACCATCTGTCTCTT 675

RESULT 12
ADL46204 standard; DNA; 678 BP.
ID      ADL46204
XX
XX      ADL46204;
XX
XX      20-MAY-2004 (first entry)
XX
XX      Discosoma red fluorescent protein (DeRed) coding sequence.
XX
XX      de; gene; red fluorescent protein; DeRed; fluorescence; red wavelength;
XX      oligomerization; tetramerization; immunoassay; hybridization assay.
XX
XX      Discosoma sp.
XX
XX      Key      Location/Qualifiers
XX      FH      1..678
XX      CDS      /*tag= a
XX      FT      /product= "DeRed protein"
XX      PN      WO2003086446-A1.
XX
XX      23-OCT-2003.
XX
XX      09-APR-2003; 2003WO-US010879.
XX
XX      10-APR-2002; 2002US-00121258.
XX      29-JUL-2002; 2002US-00209208.
XX
XX      (REGC ) UNIV CALIFORNIA.
```

XX Telen RX, Campbell RE, Baird GS;  
 XX WPI; 2003-845265/78.  
 DR P-PSDB; ADL46203.  
 XX New monomeric and dimeric Anthozoan fluorescent protein variants with  
 PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
 PT in molecular biology, e.g. in immunoassays or in tracking protein  
 PT movement in cells.  
 XX Disclosure; SEQ ID NO 2; 166pp; English.  
 PS  
 XX  
 XX The invention relates to a polynucleotide sequence encoding a Discosoma  
 CC red fluorescent protein (DERED) variant having a reduced propensity to  
 CC oligomerize. The protein variant comprises one or more amino acid  
 CC substitutions at the AB and/or AC interface(s) of the wild-type Dered  
 CC sequence, where the substitutions result in reduced propensity of the  
 CC Dered variant to form tetramers and where the variant displays detectable  
 CC fluorescence of at least one red wavelength. The composition and methods  
 CC are useful in producing red fluorescent proteins having reduced  
 CC propensity for oligomerization, especially tetramerization. The protein  
 CC may be used in molecular biology and in other scientific applications,  
 CC such as in immunoassays or hybridization assays, or in tracking the  
 CC movement of proteins in cells. This sequence corresponds to the Dered  
 CC coding sequence.

XX Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	2_3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	11	Gaps:	0

US-10-006-922a-12 (1-225) x ADL46204 (1-678)

QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20  
 DB 1 ATAGGCTTCCTCAAGAAATGTTATCAAGAGATTCATGAGGTTTAAAGTTCCCATGGAAGGA 60  
 QY 21 ThrValAsnGlyHISGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
 DB 61 ACCGTCATATGGGACGACGATTGAAATAGAAAGGAGAGAGGAGGCCATACGAAAGGC 120  
 QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTPASPile 60  
 DB 121 CACAAATCCGTAAGCTTAAGGTAAACCAAGGGGAGACCTTGGCCATTGGCTGGGATATT 180  
 QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
 DB 181 TTGTACACCAAAATTCAGTATGAGCAAGGATATATGTCAGACCCCTGCCGACATACCA 240  
 QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTyrGluArgValIleAsnPheGlu 100  
 DB 241 GACTATTAATAAGCTCATTTCCGAAAGGATTTAAATGGGAAAGGTCATGAATTTGAA 300  
 QY 101 AspGlyIysValValIleThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 DB 301 GACGCTGGCTCGTACTGTAACCCAGAGATTCAGTTGGCAGATGGCTGTTTCATCTAC 360  
 QY 121 LysValIysPheIleGlyValIleAsnPheProSerAspGlyProValIleGlnIysIysThr 140  
 DB 361 AAGGTCAAGTTCATGGCGTGAACCTTCTCCATGGAACCTGTTATGCAAAAGAAAGCA 420  
 QY 141 MetGlyTyrGluIleAsnThrGluArgLeuTyrProArgAspGlyValIleIysGlyGlu 160  
 DB 421 ATGGGCTGGGAGGACGACGACGATTCGTTGATCTCGTATGCGGCTGTTAAAGGAGAG 480  
 QY 161 IleHisIysAlaLeuIysLeuIysAspGlyIysHisTyrLeuValGluPheIysSerIle 180

DB 481 ATTCAATTAAGGCTCTGAAGCTGAAGACGCTGGTCATTAACCTAGTGAATCAAAAGTATT 540  
 QY 181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrValIleAspSerIysLeuAsp 200  
 DB 541 TACATGGCAAAAGACCTGTGACGATACCAAGGCTACTACTATGATGATCCAAACGAGAT 600  
 QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
 DB 601 ATPAACAGCCACACGAAACATATACATGCTTGAAGCACTATGAAAGAACCAAGGAGCCG 660  
 QY 221 HisIleLeuPheLeu 225  
 DB 661 CACCACTGTTCTTCTT 675  
 RESULT 13  
 ADN33978  
 ID ADN33978 standard; DNA; 678 BP.  
 XX  
 AC ADN33978;  
 XX  
 DT 01-JUL-2004 (first entry)  
 XX  
 DB Wild-type Dered encoding sequence.  
 XX  
 KW Chidarian; fluorescence resonance energy transfer; FRET; wild-type Dered;  
 KW db.  
 OS Discosoma sp.  
 XX  
 OS  
 PN WO2003054158-A2.  
 XX  
 PD 03-JUL-2003.  
 XX  
 PF 18-DEC-2002; 2002WO-US040539.  
 XX  
 PR 19-DEC-2001; 2001US-0341723P.  
 XX  
 PA (UYCH-) UNIV CHICAGO.  
 XX  
 PI Bevis B, Glick B;  
 XX  
 DR WPI; 2003-569236/53.  
 DR P-PSDB; ADN33979.  
 XX  
 PT Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent  
 PT mutant of a Chidarian chromo- or fluorescent protein or its mutant,  
 PT useful for applications involving chromo- or fluorescent proteins.  
 XX  
 PS Claim 8; SEQ ID NO 1; 65pp; English.  
 XX  
 XX The present invention relates to nucleic acid that encodes a rapidly  
 CC maturing chromo or fluorescent mutant of a Chidarian chromo- or  
 CC fluorescent protein or its mutant. The protein is useful in applications  
 CC involving nucleic acid encoding a chromo- or fluorescent protein and is  
 CC useful for producing a chromo and/or fluorescent protein which involves  
 CC growing the cell, whereby the protein is expressed, and isolating the  
 CC protein substantially free of other proteins. The protein is useful in  
 CC applications involving chromo- or fluorescent protein and is useful as  
 CC PCR primers, hybridization probes, etc. The expression cassettes are  
 CC useful for synthesizing related proteins. The chromoproteins are useful  
 CC as coloring agents which are capable of imparting color or pigment to a  
 CC particular composition of matter e.g. food compositions, pharmaceuticals,  
 CC cosmetics, living organisms, e.g., animals and plants. The chromoproteins  
 CC may also find use as labels in analyte detection assays, e.g. assays for  
 CC biological analytes of interest and as selectable markers in recombinant  
 CC DNA applications, e.g. the production of transgenic cells and organisms.  
 CC The fluorescent proteins find use in a variety of different applications,  
 CC e.g. in fluorescence resonance energy transfer (FRET) applications, as  
 CC biosensors in prokaryotic and eukaryotic cells, in applications involving  
 CC the automated screening of arrays of cells expressing fluorescent  
 CC reporting groups by using microscopic imaging and electronic analysis, as  
 CC second messenger detectors, and in fluorescence activated cell sorting  
 CC applications and as in vivo marker in animals. The fluorescent proteins

CC also find use in protease cleavage assays. The proteins can also be used  
CC is assays to determine the phospholipid composition in biological  
CC membranes and as a fluorescent timer. The present sequence represents the  
CC wild-type DsRed encoding sequence.

XX Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Score: 2,3e-140 Length: 678  
Percent Similarity: 1214.00 Matches: 225  
Best Local Similarity: 100.00% Conservative: 0  
Query Match: 100.00% Mismatches: 0  
DB: 11 Indels: 0  
Gaps: 0

US-10-006-922A-12 (1-225) x ADN33978 (1-678)

QY 1 MetATGSeSerLyAsnValIleLySGluphMetArGpHeLyValArgMetGluGly 20  
DB 1 ATGAGGTCCTCCAAAGATTATTCAGAGAGCTTCATGAGTTTAAGTTTCGATGGAAGA 60  
QY 21 ThrValaenGlyVHieGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40  
DB 61 ACGGTCATAGGGCAAGATTGAAATAGAAAGCGAAGGAGGAGCGCATACGAAGGC 120  
QY 41 HisAenThrValLyLeuLyValThrlYsGlyGlyProLeuProPheLaATrAaplle 60  
DB 121 CACCAATACCGTAAAGCTTAAGTTAACCAAGGGGAGCTTTGCCATTTCGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrglySerLyValTyValLySHieProLaAapllePro 80  
DB 181 TTGTACCAACCAATTTCAATATGGAAGCAAGTATATGTCAAGCAACCTTCGACATACCA 240  
QY 81 AppTyryLyLeuLeuSerPheProGluGlyPheLySTyrglyArgValMetAenPheGlu 100  
DB 241 GACTATAAAGCTGCTCATTTCCGTGAAGGATTTAAATGGAAAGGGTCATGAACCTTGAA 300  
QY 101 AppGlyLyValValThrlValThrglnAapSerSerLeuGlnAapGlyVcyPheHeIleTy 120  
DB 301 GACGGTGGGGGTGATCTGTAAACCCAGAGTCCAGTTTCAGAGTGGCTGTTTCATCTAC 360  
QY 121 LySValLyPheIleGlyValAenPheProSerAapGlyProValMetGlnLySlyThr 140  
DB 361 AAGGTCATGCTCATTTGCGGTGAACCTTCTTCGATGAGACCTGTATGCAAAAGAAAGA 420  
QY 141 MetGlyTyrglyAlaSerThrglyArgLeuTyProArgAapGlyValLeuLySlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGATGAGCGCTGTTGAAAGGAGAG 480  
QY 161 IleHieLyValLeuLyLeuLyAapGlyVcyVHieTyLeuValGluPheLySserIle 180  
DB 481 ATTCAATTAAGGCTCTGAAGCTGAAGACGGGTGCTCATTTACCTAGTTGAATTCAAAGTATT 540  
QY 181 TyrMetAlaLyLeuLyProValGlnLeuProGlyTyTyTyValAapSerLySleuAap 200  
DB 541 TACATGCGCAAAAGGCTGTGACGCTACCAAGGTACTACTATGTTGACTCCAAACGTGAT 600  
QY 201 IleThrSerHieAenGluAapTyThrlIleValGluGlnTyrglyArgThrgluGlyArg 220  
DB 601 ATTAACAAGCACAAACGAAGACTATCAATCGTTGAGAGTAGAAGAACCGAGGGAGCGC 660  
QY 221 HisHieLyLeuPheLeu 225  
DB 661 CACCAATCTGTCTCTT 675

RESULT 14  
AD136420  
ID AD136420 standard; DNA; 678 BP.  
AC AD136420;  
XX  
XX 22-APR-2004 (first entry)  
DT  
XX

DE Discosoma sp. red fluorescent protein (RED) DNA.

XX Fluorobody; binding ligand; green fluorescent protein; GFP;  
KW target detection; gene; red fluorescent protein; ds; RED.  
XX

OS Discosoma sp.

XX Key Location/Qualifiers  
FT CDS 1..678  
FT /tag= a  
FT /product= "Discosoma sp. red fluorescent protein (RED)"

US2003203355-A1.

30-OCT-2003.

24-APR-2002; 2002US-00132067.

24-APR-2002; 2002US-00132067.

(LALA-) LOS ALAMOS NAT LAB.  
(RBCG) UNIV CALIFORNIA.

Bradbury AM, Zeytun A, Waldo GS;

MP1: 2004-154325/15.  
P-PSDB; AD136421.

PT Novel binding ligand with intrinsic fluorescence and comprising  
PT fluorescent protein having heterologous binding sites, useful for  
PT detecting target molecule.  
XX

PS Example 6; SEQ ID NO 3; 23pp; English.

CC The invention relates to binding ligands (fluorobodies) with intrinsic  
CC fluorescence, which comprises green fluorescent protein (GFP) having  
CC heterologous binding sites. The binding ligand is useful for detecting  
CC the target molecule and is efficiently detects the target molecule. The  
CC present sequence is Discosoma sp. red fluorescent protein (RED) DNA used  
CC in the exemplification of the invention.

XX Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-10-006-922A-12 (1-225) x AD136420 (1-678)

QY 1 MetATGSeSerLyAsnValIleLySGluphMetArGpHeLyValArgMetGluGly 20  
DB 1 ATGAGGTCCTCCAAAGATTATTCAGAGAGCTTCATGAGTTTAAGTTTCGATGGAAGA 60  
QY 21 ThrValaenGlyVHieGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40  
DB 61 ACGGTCATAGGGCAAGATTGAAATAGAAAGCGAAGGAGGAGGCGCATACGAAGGC 120  
QY 41 HisAenThrValLyLeuLyValThrlYsGlyGlyProLeuProPheLaATrAaplle 60  
DB 121 CACCAATACCGTAAAGCTTAAGTTAACCAAGGGGAGCTTTGCCATTTCGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrglySerLyValTyValLySHieProLaAapllePro 80  
DB 181 TTGTACCAACCAATTTCAATATGGAAGCAAGTATATGTCAAGCAACCTTCGACATACCA 240  
QY 81 AppTyryLyLeuLeuSerPheProGluGlyPheLySTyrglyArgValMetAenPheGlu 100  
DB 241 GACTATAAAGCTGCTCATTTCCGTGAAGGATTTAAATGGAAAGGGTCATGAACCTTGAA 300

QY 101 AspGlyValValAlaThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGCTCGTACTGTAACCCAGATTCCAGTTGCGATGGCTCTTCATCTAC 360  
QY 121 LysValIysPheIleGlyValIAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGTCAAAGTTCATTTGGCGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCTGTATGGCGCTTTGAAAGAGAG 480  
QY 161 ILeHisIysAlaLeuLysLeuLysAspGlyValHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATTCATTAAGCTCTGAAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValIAspSerLysLeuAsp 200  
DB 541 TACATGGCAAGAGAGCTGTGACACTACAGGGGTACTATATGTTGACTCCAAACTGGAT 600  
QY 201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATTAACAAGCCACACAGAACTATACATCGTTAGACATATGAAGAAGCAGAGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCCCTT 675  
RESULT 15  
ADM97768 standard; DNA; 678 BP.  
XX ADM97768;  
AC ADM97768;  
XX 01-JUN-2004 (first entry)  
XX D sp red fluorescent protein coding sequence SEQ ID NO: 21.  
XX day, enzyme; sensor cell; fluorescent protein;  
XX signal transduction detection system; promoter; targeting sequence;  
XX targeted drug.  
XX Diacosome sp.  
XX Key Location/Qualifiers  
FT 1..678  
FT /\*tag= a  
FT /product= "fluorescent protein"  
XX WO2004031415-A2.  
XX 15-APR-2004.  
XX 05-SEP-2003; 2003WO-US028078.  
XX 05-SEP-2002; 2002US-0408297P.  
XX (VERT-) VERTEX PHARM INC.  
XX Whitney MA, Zeh K, Sanders PS;  
XX WPI, 2004-330208/30.  
XX P-P8DB; ADM97769.  
XX Developing a sensor cell, useful in determining the activity of a target  
XX gene and in developing therapeutic drugs, comprises providing cells  
XX comprising a signal transduction detection system and introducing DNA  
XX construct into cells.  
XX Disclosure; Page 167-168; 234pp; English.  
XX The present invention relates to a method of developing a sensor cell,  
XX for determining the activity of a target gene in the cell, which

CC comprises providing a homogeneous population of cells, where each of the  
CC cells comprises a signal transduction detection system and introducing  
CC into the population of cells an isolated DNA construct comprising a  
CC promoter operatively linked to a targeting sequence. The method is useful  
CC in developing a sensor cell for determining the activity of a target gene  
CC in the cell. The sensor cell and the methods are useful in developing new  
CC and therapeutic drugs directed to the targets. The present sequence is a  
CC coding sequence shown in the exemplification of the invention.  
XX  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2,3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0  
US-10-006-922a-12 (1-225) x ADM97768 (1-678)  
QY 1 MetArgSerSerIysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGAGGTCTTCCAAGATGTTATCAAGAGATTCAAGGTTTAAGGTTTCGATGGAAGA 60  
QY 21 ThrValAsnGlyHisGluPheGlnIleGluGlyValGluArgProTyrGluGly 40  
DB 61 ACGGTCAATGGCGACGAGTTGAAATAGAAAGCCAAAGAGAGAGGAGCCATACAAAGCC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyValProLeuProPheAlaTyrAspIle 60  
DB 121 CACAAATCCGTAAGCTTAAAGTACCAAGGGGAGACTTTGCCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro 80  
DB 181 TTGTCAACAAGATTTCAGTATGAGAAAGCATATATGCAAGCACCTGCGCAATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrProGluArgValMetAsnPheGlu 100  
DB 241 GACTATTAAGAAAGCTGATCTGAAAGATTTAAATGGAAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyValValAlaThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTGTAACCTGTAACCCAGATTCCAGTTGCAAGATGGCTGTTTATCTAC 360  
QY 121 LysValIysPheIleGlyValIAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGTCAAAGTTCATTTGGCGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCTGTATGGCGCTTTGAAAGAGAG 480  
QY 161 ILeHisIysAlaLeuLysLeuLysAspGlyValHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATTCATTAAGCTCTGAAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValIAspSerLysLeuAsp 200  
DB 541 TACATGGCAAGAGAGCTGTGACACTACAGGGGTACTATATGTTGACTCCAAACTGGAT 600  
QY 201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATTAACAAGCCACACAGAACTATACATCGTTAGACATATGAAGAAGCAGAGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCCCTT 675  
RESULT 16  
AAA48743 standard; cDNA; 695 BP.  
XX AAA48743  
XX

AC AAA48743;  
 XX 19-SEP-2000 (first entry)  
 DT Humanised Discosoma sp. "red" novel fluorescent protein drfp583 cDNA.  
 XX DE Anthozoa; drfp583; fluorescent protein; non-bioluminescent organism;  
 XX KM fluorescent labelling; ss.  
 XX OS Discosoma sp.; "red".  
 XX OS Synthetic.  
 XX PN MO200034326-A1.  
 XX PD 15-JUN-2000.  
 XX PF 10-DEC-1999; 99MO-US029473.  
 XX PR 11-DEC-1998; 98US-00210330.  
 XX PR 14-OCT-1999; 99US-00418529.  
 XX PA (CLON-) CLONTECH LAB INC.  
 XX PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
 PI Ding L;  
 XX MPI; 2000-42381/36.  
 XX DR Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
 PT useful for fluorescent labeling and as markers.  
 XX Claim 6; Page 75-76; 86pp; English.  
 XX SQ The present sequence is humanised drfp583 cDNA. drfp583 is a full-length  
 CC cDNA encoding a novel fluorescent protein (nfp) from Discosoma sp. "red",  
 CC a non-bioluminescent species of the Class Anthozoa. The wild-type drfp583  
 CC nucleotide sequence was altered to optimise the codons for expression of  
 CC the fluorescent protein in mammalian cells. Fluorescent proteins can be  
 CC used in fluorescent labeling, a useful tool for marking a protein, cell  
 CC or organism of interest. Unlike other markers used in protein labeling,  
 CC such as beta-galactosidase and luciferase, fluorescent proteins do not  
 CC require an exogenous cofactor or substrate. Methods involving fluorescent  
 CC proteins are also less laborious and less difficult to control than the  
 CC traditional methods of fluorescent labeling, where a protein of interest  
 CC is purified and then covalently conjugated to a fluorophore derivative.  
 CC Novel fluorescent proteins isolated from species of the Class Anthozoa  
 CC can be used as markers for gene expression and protein localization  
 CC studies, and in fluorescence resonance energy transfer (FRET) reactions.  
 CC They may have improved properties and better suitability for larger  
 CC excitations compared to prior art fluorescent proteins such as green  
 CC fluorescent protein  
 XX SQ Sequence 695 BP; 149 A; 228 C; 209 G; 109 T; 0 U; 0 Other;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 2,39e-140 Length: 695  
 XX Score: 1214.00 Matches: 225  
 XX Percent Similarity: 100.00% Conservative: 0  
 XX Best Local Similarity: 100.00% Mismatches: 0  
 XX Query Match: 100.00% Indels: 0  
 XX DB: 3 Gaps: 0  
 XX US-10-006-922a-12 (1-225) x AAA48743 (1-695)  
 QY 1 MetArgSerSerIysAsnValIleIyGluPheMetArgPheIysValIArgMetGluGly 20  
 DB 10 ATCGGCTCCCTCAAGAACGTCAACAGAGATTCAATCGGCTTCAAGGCGCCAGAGAGGC 69  
 QY 21 ThrValAsnGlyIhIegIupheGluIleGluGlyGluGlyIArgProTyrgIuGly 40  
 DB 70 ACCGTGAACGCGCAGAGGTTCAAGATCAAGGCGAGGCGAGGCGCCCTACGAGGCG 129  
 QY 41 HisAsnThrValysLeuIysValThrIyGlyGlyProIeuProPheIaIATpApIle 60

DB 130 CACACACCCGTGAAGCTGAAGGTGACCAAGGCGGCCCCCTGCGCTTCCGCTGGACATC 189  
 QY 61 LeuSerProGlnPheGlnTyrgIySerIysValTyValIyshiProAlaapIlePro 80  
 DB 190 CTGTCCCCCAGATTCCAGTAGGCTCCAGAGGTGTACGTGAAGACCCCCCGGCATCCCC 249  
 QY 81 AapTyrgIyIyLeuSerPheProGluGlyPheIyVtTgIuArgValMetAsnPheGlu 100  
 DB 250 GACTACAAAGAGCTGCTTCCCGAGGCTTCAGTGAAGCGCGGTGAATGAATCTTCGAG 309  
 QY 101 AapGlyGlyValValThrValThrGlnApeSerSerLeuGlnAapGlyCyshIleTyT 120  
 DB 310 GACGCGGCGGTGTGACCGGTGACCCAGACACTCTCCCTCGACGAGCGGCTTCATCTAC 369  
 QY 121 IyValIyPheIleGlyValIAsnPheProSerAapGlyProValMetGluIyVtT 140  
 DB 370 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGAGCGCCGTGATGCAAGAAAGACC 429  
 QY 141 MetGlyTTrGluIAserThrGluArgLeuTyTProArgAapGlyValIeIyVtT 160  
 DB 430 ATGGGCTGGAGGCTTCACCGAGCGCTGTACCCCGGACGCGGTGTGAAGGCGAG 489  
 QY 161 ILehIyAlaIeIyIshIyAapGlyGlyIyIshIyTyrIeIyValGluPheIySserIle 180  
 DB 490 ATCCACAAAGGCCCTGAAGCTGAAGAGACGCGCGCACTACTGTGTGAAGTTCATC 549  
 QY 181 TyrMetAlaIyIyAapProValGluIeuProGlyTyTyrTyValIAsPserIyLeuAap 200  
 DB 550 TACATGGCCAAAGAGCCCGTGCAGCTGCGCGGTACTACTGTGACTCCAAAGCTGGAC 609  
 QY 201 ILehSerHisAsnGluAapTyTTrIleValGluGlnTyTgIuArgThrGluGlyArg 220  
 DB 610 ATCACTCCCAACAGAGACTACACATCGTAGAGAGTACAGAGCGACAGAGGCGGC 669  
 QY 221 HLehIeIyPheIeu 225  
 DB 670 CACCACTGTTCCTG 684  
 RESULT 17  
 AAL47952  
 ID AAL47952 standard; DNA; 859 BP.  
 XX AAL47952;  
 AC 26-SEP-2002 (first entry)  
 DT  
 XX 26-SEP-2002 (first entry)  
 XX DE Discosoma red fluorescent protein coding sequence.  
 XX KM Yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette;  
 XX KM modified yeast strain; environmental pollution; gene; ds.  
 XX OS Discosoma sp.  
 XX FT Key Location/Qualifiers  
 FT CDS 54..731  
 FT /\*tag= a  
 FT /product= "red fluorescent protein"  
 PN DE10061872-A1.  
 XX PD 20-JUN-2002.  
 XX PF 12-DEC-2000; 2000DB-01061872.  
 XX PR 12-DEC-2000; 2000DB-01061872.  
 XX PA (LICH/) LICHTENBERG-FRATZ H.  
 XX PI Lichtenberg-Frate H;  
 XX DR MPI; 2002-539633/58.  
 DR P-FSDB; AAO18270.



XX Modified yeast strain, useful for detecting toxic compounds in  
 PT environment, contains integrated cassette responsive to genotoxic and  
 PT cytotoxic compounds.

XX Disclosure, Page 20-21; 34pp; German.

XX The present invention relates to a modified yeast strain that contains,  
 CC integrated stably and functionally in its genome, a genotoxicity cassette  
 CC and a cytotoxicity cassette, each comprising a promoter and reporter  
 CC gene, both of which are different in the two cassettes. The modified  
 CC yeast strain is used to detect environmental pollution, especially  
 CC genotoxic and/or cytotoxic substances in complex environmental  
 CC contaminants, especially organic compounds, but also (non-)ionizing  
 CC radiation and chemical carcinogens. Particular applications are in  
 CC monitoring (waste) water (e.g. as an early warning system), medical  
 CC toxicology screening and for industrial process control. The present  
 CC invention is a marker gene suitable for use in the cassettes of the  
 CC present invention

XX Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.22e-140 Length: 859  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0

US-10-006-922a-12 (1-225) x AAL47952 (1-859)

QY 1 MearGSeSerLybAaenValIleGluPhemeArGpHeLyValArGMeGluGly 20  
 Db 54 ATGAGGCTTCCAGATGTTATCAAGAGTTCATGAGTTTAAAGTTCCGATGGAAGA 113  
 QY 21 ThValAaenGlyHAgLuPhenGluIleGluGlyGluGlyArGProTyGluGly 40  
 Db 114 ACGGTCAATGGGACGAGTTGAAATAGAGCGAAGAGGAGGAGCCATACGAAGGC 173  
 QY 41 HlaaenThValLybLeuLybValThThyGlyGlyProLeuProPhaIaTTpAaPle 60  
 Db 174 CACAAATCCGTAAAGCTTAAAGTAAACAAAGGAGGAGCCCTTGGCTTGGATAT 233  
 QY 61 LeuSerProGluPhenGluTyGlySerLybValTyValLybSHaSPoAlAaPlePro 80  
 Db 234 TTGTCAACCAATTTCAATATGAGAGCAAGATATATGTCAGACACCTCCGACATACCA 293  
 QY 81 AaPtyLybLybLeuSerPheProGluGlyPheLybTyGluArGValMeLAsnPheGlu 100  
 Db 294 GACATTAATAAAGCGTCACTTCCTGAGAGATTAAATGGAAGGTCATGAACTTTGAA 353  
 QY 101 AaPtyLybValValThValThThGlnAaSPeSerLeuGlnAaPtyCybPheIleTy 120  
 Db 354 GACCGTGCGTCTGTAATGTAACCAAGATTCACATTTGACAGAGTGTTCATCTAC 413  
 QY 121 LybValLybPheIleGlyValAaSPeProSerAaPtyProValMeGluLybLybTh 140  
 Db 414 AAGGTCAAGTTCATTTGGCGTGAACCTTCCCATGAGCCTGTTATGCAAAAGAACCA 473  
 QY 141 MeGlyTyGpGluAaSerThGluArGLeuTyProArGAsPtyValleuLybGlyGlu 160  
 Db 474 ATGGGCTGGGAAGCAGCAGCAGCGTTTGTATCTCGATAGGCGGTGTGAAAGAGAG 533  
 QY 161 HlaaenLybAlaLeuLybLeuLybAaPtyGlySHaSTyLeuValGluPhelySerIle 180  
 Db 534 ATTCATAAGGCTCTAAGCTGAAGACGATGATCTTACCTAAGTGAATTCAAAAGTAT 593  
 QY 181 TyrMeAlaLybLybProValGluLeuProGlyTyTyTyTyValAaSPeSerLybLeuAa 200  
 Db 594 TACATGGCAAGAACCTGTGCACTACCAAGGATCTACTATGTTGATCCCAACCTGAT 653  
 QY 201 IleThSerHlaaenGluAaPtyThThIleValGluGlnTyGluArGThrcGluGlyArG 220

Db 654 ATTAACAGCCACAAAGACATTAACATGTTGACGATATTAAGAACCGAGAGCG 713  
 QY 221 HlaaenPheLeu 225  
 Db 714 CACCATCTGTTCCCT 728

RESULT 18

ADY51715

ID ADY51715 standard; DNA; 859 BP.

XX AC ADY51715;

DT 05-MAY-2005 (first entry)

XX Discosoma sp. red fluorescent protein (Rfp) DNA Seq 11.

XX fluorescence; mutagenesis; red fluorescent protein; gene; ds;  
 KW protein interaction.

XX Discosoma sp.

XX Key Location/Qualifiers

FT CDS 54..731  
 FT /tag= a  
 FT /product= "RFP protein"

XX WO200268605-A2.

XX 06-SEP-2002.

XX 26-FEB-2002; 2002WO-US006063.

XX 26-FEB-2001; 2001US-00794308.

XX 24-MAY-2001; 2001US-0086538.

XX (REGC ) UNIV CALIFORNIA.

XX Tsien RY, Baird GS, Campbell RE, Zacharias DA;

XX WPI; 2002-713372/77.

XX P-PSDB; ADY51716.

XX New non-oligomerizing fluorescent protein containing at least one  
 PT mutation that reduces or eliminates the ability of the protein to  
 PT oligomerize, useful for making better and new assays for molecular  
 PT biology.

XX Disclosure; SEQ ID NO 11; 117pp; English.

XX This invention relates to a novel non-oligomerizing fluorescent protein.  
 CC Specifically, it refers to the presence of at least one mutation in the  
 CC fluorescent protein that reduces or eliminates the ability of the protein  
 CC to oligomerize. The present invention describes fluorescent proteins and  
 CC derived from naturally occurring green or red fluorescent proteins and  
 CC provides a fusion protein that comprises a non-oligomerizing fluorescent  
 CC protein linked to at least one protein of interest. As such, these fusion  
 CC proteins can be used in methods and compositions to determine the pH of a  
 CC sample, or whether the sample contains an enzyme, molecule or agent that  
 CC regulates the activity of an expression control sequence. Furthermore,  
 CC they may be used to identify a specific interaction of molecules, such  
 CC that they are useful for improving or developing new assays in the field  
 CC of molecular biology. This polynucleotide is the DNA sequence that  
 CC encodes the Discosoma sp. red fluorescent protein (RFP) of the invention.

XX Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.22e-140 Length: 859  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0

```
DB: 7 Gaps: 0
US-10-006-922a-12 (1-225) x ADY51715 (1-859)
QY 1 MetArgSerSerlybAnvAlIlelysgluPhemeArgPheIyValaArgMetGluGly 20
Db 54 ATGAGGCTCTCCAAAGATGTTATCAAGAGTTCATGAGGTTTAAGGTTCCATGAAAGG 113
QY 21 ThrValaAngIyHiegluPhegluIlegluGlygluGlyIyArgProTyrgluGly 40
Db 114 ACCGTCAATGGGCACGAGTTTGAATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 173
QY 41 HisAenThrValIlyleuIyValIThrlysgIyGlyProleuProPheIaTPaapIle 60
Db 174 CACAAATACCTTAAGCTTAAGGTAAACCAAGGGGAGCTTGGCATTTGCTGGGATATT 223
QY 61 LeuSerProGlnPheGlnTyrgIySerlyValTyValIyshiAProAlaapIlePro 80
Db 234 TTGTCAACCAATTTGATGATGAGCAAGTATATGTCAAGCAACCTGCCGACATACCA 293
QY 81 AspTyrllyIyLeuSerPheProGluGlyPheIyTfPgluArgValMetAenPheGlu 100
Db 294 GACTATATAAAAGCTGCATTTCTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 353
QY 101 AspGlyGlyValIyValIThrValIThrGlnAapSerSerleuGlnAapGlyCyPheIleTy 120
Db 354 GACGGTGGCGTCTGATCTGTAACCCAGATTCACAGTTTGCAAGATGGCTTTTCATCTAC 413
QY 121 LyseValIyPheIleGlyValIaenPheProSerAapGlyProValMetGlnIylyeThr 140
Db 414 AAGGTCAAGTTCAATTGGCGTGAACCTTCTCCGATGACCTGTATGCAAAAGAGACA 473
QY 141 MetGlyTfPgluAlaSerThrGluArgleuTyProArgAapGlyValIleuIysgIyGlu 160
Db 474 ATGGGCTGGGAACCGACCTGACGCTGATTCCTCGTATGCGCGTTGAAAGAGAG 533
QY 161 IlleHisIyValaIeulysleuIyAapGlyGlyHieTyrlleuValIgluPheIySerIle 180
Db 534 ATTCATTAAGCTCTGAAAGCTGAAAGACGGTGGTCATTTACCTTAAGTTGAATTCAAAAGATTT 593
QY 181 TyrMetAlaIylybProValGlnLeuProGlyTyTyTyValIaapSerIylyleuAap 200
Db 594 TACATGGCAAAAGAGCTGTGACGTACCAAGGATCACTATGTGATCCAAACCTGGAT 653
QY 201 IlleThSerIleAngIyAapTyTyThrIleValIgluGlnIyTyrgIuArgThrGluIyArg 220
Db 654 ATTAACAAGCCACAAGAACTATACATCGTTGAGCATATGAAAGAACCGAGGAGCGC 713
QY 221 HisHisIeupheIeu 225
Db 714 CACCATCTGTTCCTT 728
RESULT 19
AADS3432
AADS3432 standard; DNA; 859 BP.
AC AADS3432;
XX 28-MAY-2003 (first entry)
DE Discosoma species red fluorescent protein (RFP) encoding DNA.
XX
XX Phosphorylation indicator; fluorescent protein; detection; phosphatase;
XX kinase; red fluorescent protein; RFP; gene; db.
XX
XX Discosoma sp.
XX
XX Key Location/Qualifiers
XX CDS 54..731
XX FT /tag= a
XX FT /product= "Discosoma sp. red fluorescent protein (RFP)"
XX
XX .FN MO200295058-A2.
```

```
XX 28-NOV-2002.
PD 24-MAY-2002; 2002MO-US016955.
PF 24-MAY-2001; 2001US-00865291.
PR 24-MAY-2001; 2001US-00865291.
PA (REGC ) UNIV CALIFORNIA.
XX Tsien XY, Ting AY, Zhang J;
XX MPI; 2003-148474/14.
DR P-PSDB; AAB34962.
XX
PT Novel chimeric phosphorylation indicators, useful for detecting
PT kinase/phosphatase in samples, has donor molecule, phosphorylatable
PT domain, phosphoaminoacid binding domain, and acceptor molecule, in
PT operative linkage.
XX
PS Disclosure; Col 64-65; 38pp; English.
XX
CC The present invention relates to chimeric phosphorylation indicators
CC comprising a phosphorylation polypeptide and a fluorescent protein or in
CC operative linkage, a donor molecule, a phosphorylatable domain, a
CC phosphoaminoacid binding domain (PABD) and an acceptor molecule. The
CC phosphorylation indicators of the invention are useful for detecting
CC kinases or phosphatases in a biological sample. They are also useful in
CC high throughput analysis e.g. for detecting a kinase inhibitor or
CC phosphatase inhibitor. The present sequence is Discosoma species red
CC fluorescent protein (RFP) encoding DNA used in the invention
XX
SQ Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3,22e-140 Length: 859
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-006-922a-12 (1-225) x AADS3432 (1-859)
QY 1 MetArgSerSerlybAnvAlIlelysgluPhemeArgPheIyValaArgMetGluGly 20
Db 54 ATGAGGCTCTCCAAAGATGTTATCAAGAGTTCATGAGGTTTAAGGTTCCATGAAAGG 113
QY 21 ThrValaAngIyHiegluPhegluIlegluGlygluGlyIyArgProTyrgluGly 40
Db 114 ACCGTCAATGGGCACGAGTTTGAATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 173
QY 41 HisAenThrValIlyleuIyValIThrlysgIyGlyProleuProPheIaTPaapIle 60
Db 174 CACAAATACCGTAAAGCTTAAGGTAAACCAAGGGGAGCTTGGCATTTGCTGGGATATT 223
QY 61 LeuSerProGlnPheGlnTyrgIySerlyValTyValIyshiAProAlaapIlePro 80
Db 234 TTGTCAACCAATTTGATGATGAGCAAGTATATGTCAAGCAACCTGCCGACATACCA 293
QY 81 AspTyrllyIyLeuSerPheProGluGlyPheIyTfPgluArgValMetAenPheGlu 100
Db 294 GACTATATAAAAGCTGCATTTCTCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 353
QY 101 AspGlyGlyValIyValIThrValIThrGlnAapSerSerleuGlnAapGlyCyPheIleTy 120
Db 354 GACGGTGGCGTCTGATCTGTAACCCAGATTCACAGTTTGCAAGATGGCTTTTCATCTAC 413
QY 121 LyseValIyPheIleGlyValIaenPheProSerAapGlyProValMetGlnIylyeThr 140
Db 414 AAGGTCAAGTTCAATTGGCGTGAACCTTCTCCGATGACCTGTATGCAAAAGAGACA 473
QY 141 MetGlyTfPgluAlaSerThrGluArgleuTyProArgAapGlyValIleuIysgIyGlu 160
```

```

Db      474 ATGGCGTGGAGCCAGCACTGACGCTTTGATCTCTGTATGCGCTTGAAAGAGAG 533
Qy      161 ILeHsYsAlaLeuLYsLeuLYsAspGlyGlyHsTYrLeuValGluPhelYsSerIle 180
Db      534 ATTCAATAGAGCTCGAAGCTGGAAGACGGTGTCTATTACCTAGTTGAATCAAAAGTATT 593
Qy      181 TyMeAlaLYsLYsPProValGlnLeuProGlyTYrTYrTYrValAspSerLYsLeuAsp 200
Db      594 TACATGGCAAGAACCTGTGCACTACCAAGGTACTACTATGTGTGACTCCAACTGGAT 653
Qy      201 ILeHsSerHsAsnGluAspTYrThrIleValGlnGlnTYrGluArgThrGluGlyArg 220
Db      654 ATAAACAAGCCCAACAAGAACTATACATCGTTAGCAGTATGAAAGAACCGAGGAGCG 713
Qy      221 HIsHsLeuPhelLeu 225
Db      714 CACCATCTGTTCTT 728

RESULT 20
ID      AAD61969 standard; cDNA; 859 BP.
XX      AAD61969;
AC      AAD61969;
XX      15-JAN-2004 (first entry)
DT      15-JAN-2004 (first entry)
DE      Discosoma sp. red fluorescent protein (RFP) cDNA.
KM      Fluorescent protein; resonance energy transfer; pH; detection;
XX      red fluorescent protein; RFP; gene; ss.
OS      Discosoma sp.
XX      Location/Qualifiers
FH      Key 54..731
FT      CDS /feature a
FT      /product= "Discosoma sp. red fluorescent protein"
XX      US2003170911-A1.
PN      11-SEP-2003.
XX      26-FEB-2001; 2001US-00794308.
PR      26-FEB-2001; 2001US-00794308.
XX      26-FEB-2001; 2001US-00794308.
PA      (TSIE/) TSIEH R. Y.
PA      (ZACH/) ZACHARIAS D A.
PA      (BAIR/) BAIRD G S.
XX      Tsien RY, Zacharias DA, Baird GS;
XX      WPI; 2003-802418/75.
XX      P-PSDB; ABM00918.
XX      Fluorescent proteins containing a mutation that reduces or eliminates its
XX      ability to oligomerize which gives more reliable fluorescence resonance
XX      energy transfer results and are useful to detect molecule interaction,
XX      enzyme, or sample pH.
XX      Disclosure; Page 29-30; OP; English.
XX      The invention relates to a non-oligomerizing fluorescent protein
XX      containing a mutation that reduces or eliminates its ability to
XX      oligomerize. The fluorescent protein gives more reliable fluorescence
XX      resonance energy transfer results and are useful to detect molecule
XX      interaction, enzymes, or sample pH. These are also used to identify
XX      agents or conditions that regulate expression of control sequences. The
XX      present sequence is Discosoma sp. red fluorescent protein (RFP) cDNA
XX      Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;
Alignment Scores:

```

```

Pred. No.: 3,22e-140 Length: 859
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-006-922a-12 (1-225) x AAD61969 (1-859)
Qy      1 MetArgSerSerLYsAsnValIleLYsGluPheMetArgPheLYsValArgMetGluGly 20
Db      54 ATGAGGCTTCCAAAGATGTTATCAAGAGATTCAGTGGTTTAAAGTTCCATGGAAGA 113
Qy      21 ThrValAsnGlyHsGluPheGlnIleGluGlyGluGlyGlyArgProTYrGluGly 40
Db      114 ACGGTCAATGGGCGACGAGTTGAAATGAAGGCGAAGAGAGGAGGAGCCATACGAAGGC 173
Qy      41 HIsAsnThrValLYsLeuLYsValThrLYsGlyGlyProLeuPProPheAlaTrpAspIle 60
Db      174 CACATACCGTAAAGCTTAAAGTAAAGTAAAGGAGGAGGAGCCTTGCCATTGGCTGGATATT 233
Qy      61 LeuSerProGlnPheGlnTYrGlySerLYsValTYrValLYsHsPProAlaAspIlePro 80
Db      234 TTGTCAACCAATTTCAGTATGAAAGCAAGTATATGTCAAGCACCTTCCGACATACCA 293
Qy      81 AspTYrLYsLYsLeuSerPheProGluGlyPheLYsTrpGluArgValMetAsnPhelGlu 100
Db      294 GACTATATAAAGCTGCTCATTTCTGAAGAGATTAAATGGAAAGGCTCATGAACTTTGAA 353
Qy      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTYr 120
Db      354 GACGTGGCGTCTGTTACTGTAAACCAAGATTCCAGTTTGAGAGATGGCTGTTTCATCTAC 413
Qy      121 LYsValLYsPheIleGlyValAsnPheProSerAspGlyProValMetGlnLYsLYsThr 140
Db      414 AAGGTCAAGTTCAATGGCGTGAACCTTTCCTCCATGAGACCTGTTATGCAAAAGAGACA 473
Qy      141 MetGlyTrpGluValAsnSerThrGluArgLeuTYrProArgAspGlyValIleLYsGly 160
Db      474 ATGGGCTGGGAGGACACACTGACGCTTTGATCTCTGTATGAGCGGTGTAAAGAGAG 533
Qy      161 ILeHsYsAlaLeuLYsLeuLYsAspGlyGlyHsTYrLeuValGluPhelYsSerIle 180
Db      534 ATTCAATAGAGCTCGAAGCTGGAAGACGGTGTCTATTACCTAGTTGAATCAAAAGTATT 593
Qy      181 TyMeAlaLYsLYsPProValGlnLeuProGlyTYrTYrTYrValAspSerLYsLeuAsp 200
Db      594 TACATGGCAAGAACCTGTGCACTACCAAGGTACTACTATGTGTGACTCCAAACTGGAT 653
Qy      201 ILeHsSerHsAsnGluAspTYrThrIleValGlnGlnTYrGluArgThrGluGlyArg 220
Db      654 ATAAACAAGCCCAACAAGAACTATACATCGTTAGCAGTATGAAAGAACCGAGGAGCG 713
Qy      221 HIsHsLeuPhelLeu 225
Db      714 CACCATCTGTTCTT 728

RESULT 21
ID      ADX26533 standard; DNA; 859 BP.
XX      ADX26533;
XX      ADX26533;
AC      ADX26533;
XX      21-APR-2005 (first entry)
DT      21-APR-2005 (first entry)
DE      Discosoma DsRed RFP protein coding sequence, seq id 11.
XX      Discosoma DsRed RFP protein coding sequence, seq id 11.
XX      Phosphorylation; detection; red fluorescent protein; RFP; gene; ds.
XX      Discosoma sp.
OS      Discosoma sp.
XX      Location/Qualifiers
FH      Key 54..731
FT      CDS

```

```

FT      /*tag= a
PT      /product= "DRed red fluorescent protein"
XX
XX      US2005026234-A1.
XX
XX      03-FEB-2005.
XX
XX      28-MAY-2004; 2004US-00857622.
XX
XX      31-JAN-1996; 96US-00594575.
XX      31-JAN-1997; 97US-00792553.
XX      13-SEP-1999; 99US-00396003.
XX      24-MAY-2001; 2001US-00865291.
XX
XX      (VIOL/) VIOLIN J D.
XX      (NEWT/) NEWTON A C.
XX      (TSIE/) TSIE R Y.
XX      (ZHAN/) ZHANG J.
XX
XX      VIOLIN JD, Newton AC, Tsien RY, Zhang J;
XX      WPI; 2005-141377/15.
XX      P-PSDB; ADX26534.
XX
XX      New chimeric phosphorylation indicator comprises a first fluorescent
XX      protein, phosphaminoacid binding domain with FHA2 sequence and protein
XX      kinase C-phosphorylatable domain, useful for detecting protein kinase C
XX      or phosphates.
XX
XX      Disclosure; SEQ ID NO 11; 68bp; English.
XX
XX      The invention relates to a chimeric phosphorylation indicator (CPI) (I)
XX      comprising in operative linkage, a first fluorescent protein, a
XX      phosphaminoacid binding domain with an FHA2 sequence defined in the
XX      specification, a protein kinase C (PKC)-phosphorylatable domain, and a
XX      second fluorescent protein. The first and second fluorescent proteins are
XX      chosen from green fluorescent proteins (GFPs) and red fluorescent
XX      proteins (RFPs). The first and second fluorescent proteins exhibit a
XX      detectable resonance energy transfer, when the first fluorescent protein
XX      is excited. The PKC-phosphorylatable domain and phosphaminoacid binding
XX      domain do not substantially emit light to excite the second fluorescent
XX      protein. (I) is useful for detecting a protein kinase C (PKC) or
XX      phosphates in a sample (such as biological sample e.g. cell, tissue
XX      sample, or their extracts). (I) is useful for detecting a kinase
XX      inhibitor or phosphatase inhibitor. The method of the invention is
XX      adapted to high throughput analysis. The current sequence represents the
XX      coding sequence of Discosoma DRed RFP.
XX
XX      Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 3.22e-140 Length: 859
XX      Score: 1214.00 Matches: 225
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 14 Gaps: 0
XX
XX      US-10-006-922a-12 (1-225) x ADX26533 (1-859)
XX
XX      1 MetArGSeRSeRlySAsnVAlIleYgGluPhMeArGpHeLySValArGMeGluGly 20
XX      |||||
XX      54 ATGAGGCTTCCAGAAATGTTATCAAGAGCTTCAAGGTTTAAGGTTCCGATGGAAGCA 113
XX      |||||
XX      21 ThrValaNglyNhiGluPhNeGluIleGluGlyGluGlyGlyArGProTyRgGluGly 40
XX      |||||
XX      114 ACGGTCAATGGGCAAGATTGAAATAGAGGCAAGAGGAGGAGGCCATACGAAGGC 173
XX      |||||
XX      41 HisAenThrVallySLeuLySValThlySgLyGlyProLeuPProPheAlaTPAsPllle 60
XX      |||||
XX      174 CACAAATACCGTAAAGCTTAAGGTAAACCAAGGGGAGACCTTTGCCATTGTGGGATATT 233
XX      |||||
XX      61 LeuSerProGlnPheGlnTyRgLySLeLySValTyRVallyShISProAlaAsPlllePro 80

```

```

DB      234 TTGTCACCACAATTTCAGTATGGAACCAAGTATATGTCAAGCACCTGCCGACATACCA 293
XX      |||||
XX      81 AsPTyRlySlyLeuSerPheProGluGlyPheLySTrgLulArGyValMeLanPheGlu 100
XX      |||||
XX      294 GACTATTAATAAGCTTCATTTCTCTGAAGATTTTAATGGAAAGGGTCATGAACCTTTGAA 353
XX      |||||
XX      101 AspGlyGlyValValThrValThrgInaPseSerSeLyGlnaAspGlyCySPheIleTyR 120
XX      |||||
XX      354 GACGCTGGCGCTCTTACTGTAACCCAGATTCCAGTTTGAGAGATGGCTCTTCATCTAC 413
XX      |||||
XX      121 LysVallySPhelIleGlyValaenPheProSeRAspGlyProValMeGlnlySlyThr 140
XX      |||||
XX      414 AAGTCAGATTCAATTGGCGGAACTTCTCTCGATGACACTGTATGCAAAAGAGACA 473
XX      |||||
XX      141 MetGlyTYTgLuAlaSerThrgLuAgluTyRProArGAspGlyValleuSgLyGlu 160
XX      |||||
XX      474 ATGGGCTGGGAAACCGACACTGAGCGTTGTATCCGTGATGGCGTGTGAAAGAGAG 533
XX      |||||
XX      161 IleHisLysAlaLeuLySLeuLySAspGlyGlyNhiSTyRleuValGluPhLySertIle 180
XX      |||||
XX      534 ATTCATTAAGGCTTGAAGCTGAAAGACGGTGGTCATTACTTAATTCAAAAGATT 593
XX      |||||
XX      181 TyrMetAlaLySlySProValGlnLeuProGlyTYRtyTYRValaPseRlySleuAsp 200
XX      |||||
XX      594 TACATGGCAAGAGAGCGTGGAGCTTACAGAGGTACTACTATGTCCTCAAACTGGAT 653
XX      |||||
XX      201 IleThSerHisAenGluAspTyRThrlIleValGluGlnTyRtGluArgThrgLuGlyArg 220
XX      |||||
XX      654 ATTAACAAGCCACAACAAGACTATCAATCGTTGAGCAATATGAAAGAACCGAGGAGCGC 713
XX      |||||
XX      221 HisHisLeuPheLeu 225
XX      |||||
XX      714 CACCACTGTGTTCTT 728
XX
XX      RESULT 22
XX      ID ACA62995 standard; DNA; 3311 BP.
XX      ACAC62995;
XX      ACAC62995;
XX      23-SEP-2003 (first entry)
XX
XX      Plasmid DNA containing coding sequence for RFP.
XX
XX      Fluorescently-tagged enzyme; substrate; cell population; GFP;
XX      quantification of enzymatic activity; green fluorescent protein;
XX      red fluorescent protein; RFP; enhanced green fluorescent protein; EGFP;
XX      enzymatic process; cellular enzyme; chemotherapeutic drug;
XX      multidrug resistance; MDR; de.
XX
XX      Aequorea victoria.
XX      OS Synthetic.
XX      US2003049597-A1.
XX      PD 13-MAR-2003.
XX
XX      01-MAR-2001; 2001US-00797496.
XX
XX      01-MAR-2001; 2001US-00797496.
XX
XX      01-MAR-2001; 2001US-00797496.
XX
XX      (SIMO/) SIMON S M.
XX      PA (CHEN/) CHEN Y.
XX
XX      Simon SM, Chen Y;
XX      PI
XX      DR WPI; 2003-555145/52.
XX
XX      Simultaneously quantifying in situ the relationship between an enzyme and
XX      its substrate for study of enzymatic processes at a cellular level,
XX      comprises optical measurements on cells expressing the fluorescently
XX      tagged enzyme.
XX
XX      PT

```

XX Disclosure; Page 8-9; 21pp; English.

CC The present invention relates to a method for simultaneously quantifying  
CC in situ the relationship between a fluorescently-tagged enzyme (E) and  
CC its substrate in a population of cells. A population of cells expressing  
CC (E) is created in which different cells contain different amounts of (E).  
CC The cells are incubated with a substrate, and the concentration and  
CC enzymatic activity of (E) in each cell are simultaneously quantified by  
CC optical means. The fluorescently-tagged enzyme is produced by linking the  
CC enzyme with *Aequorea victoria* green fluorescent protein (GFP), or its  
CC variants or derivatives such as red fluorescent protein (RFP) and  
CC enhanced green fluorescent protein (EGFP). The method of the invention is  
CC useful for simultaneously quantifying in situ the relationship between an  
CC enzyme and its substrate in a cell. The method is useful for the study of  
CC enzymatic processes at the cellular level, and especially for examining  
CC the interactions of cellular enzymes with chemotherapeutic drugs, e.g.  
CC for elucidating mechanisms of multidrug resistance (MDR). The present  
CC sequence represents a plasmid containing the coding sequence for RFP  
XX

SO Sequence 3311 BP; 858 A; 778 C; 860 G; 815 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2.15e-139 Length: 3311  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x ACA62995 (1-3311)

Qy 1 MearGserSerLyAsnValIlelyGluPhemeAcrPhelyValaIarGmeGluGly 20  
Db 289 ATGAGGCTTCCAGCAATGTTTTCAGAGCTTCATGAGCTTTAAGTTCCGATGGAAGA 348  
Qy 21 ThrValaengIyHleGluPhelGluGluGluGluGluGluGluGluGluGluGlu 40  
Db 349 ACGGTCATGAGGACGAGTGTGAATAGAGCGAGAGGAGGAGGAGGAGGAGGAGGAG 408  
Qy 41 HisGenthrValLyLeuLyValThrLyGlyGlyProLeuProPhalaTtrAaplle 60  
Db 409 CACATATACCGTAAGCTTAAGGTAAACAAAGGGGAGCCCTTGCCATTGCTGGATAT 468  
Qy 61 LeuSerProGlnPhelGlnTyrGlySerLyValTyrValLyHleProAlAapllePro 80  
Db 469 TTGTCCACCACTTTCACTATGGAAGCAAGGTAATGTCTACAGACCCCTCCGACATACA 528  
Qy 81 AapTyrLyLyLeuSerPheProGluGlyPheLySTrPGluArgValaIarGmeGlu 100  
Db 529 GACATATAAAGCTGTCACTTCTCTCAAGATTTAATGAGAAAGGCTCATGAACCTTGA 588  
Qy 101 AapGlyLyValaIthrValThrGlnAapSerSerLeuGlnAapGlyCyPheHleTyr 120  
Db 589 GACGGTGGCGTGTACGTAAACCAAGATTCACATTTGACAGATGGCTGTTCACTTAC 648  
Qy 121 LyValLyPheHleGlyValaAsnPheProSerAapGlyProValaIarGmeGlu 140  
Db 649 AAGGTCAAGTTCACTTGGGTGACCTTCTTCGATGACCTGTTATGCAAAAGAAACA 708  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAapGlyValaIarGmeGlu 160  
Db 709 ATGGGCTGGAAAGCAGCAGCTAGCGTTGTATCTCCGAGATGGCGTGTGAAGAAGAG 768  
Qy 161 IleHleValaIarLeuLyLeuLyAapGlyGlyHleTyrLeuValaIarGmeGlu 180  
Db 769 ATTCTTAAGGCTCTAAGCTGAAGCGGTGTCTATTCATAGTTGAATCAAAAGTAT 828  
Qy 181 TyrMetAlaLyLeuProValaGlnLeuProGlyTyrTyrTyrValaIarGmeGlu 200  
Db 829 TACATGGCAAAAGCCTGTGACGCTTACCAAGGCTACTACTATGTTGATCCAAACTGAT 888  
Qy 201 IleThrSerHleAsnGluAapTyrThrIleValaIarGlnTyrGluArgThrGluGlyArg 220

Db 889 ATTAACAGCCACACAGACGATACATGCTTGACAGATGTAAGAAACGAGGAGCC 948  
Qy 221 HisHlePheLeu 225  
Db 949 CACCATCTGTCTCT 963

#### RESULT 23

ID ADC24134 standard; DNA; 681 BP.

AC ADC24134;

DT 18-DEC-2003 (first entry)

DE Discosoma wild-type red fluorescent protein DNA #2.

KM Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
KM fluorescent protein variant; transcription induction detection;  
KM fluorescence energy resonance transfer; FRRET; protein kinase;  
KM protein phosphatase; ion indicator; ds.

OS Discosoma.

XX US2003059835-A1.

XX 27-MAR-2003.

XX 10-APR-2002; 2002US-00121258.

XX 26-FEB-2001; 2001US-00794308.

XX 24-MAY-2001; 2001US-0086538.

XX (TSIE//) TSIE R Y.

XX (CAMP//) CAMPBELL R E.

XX Tsien RY, Campbell RE;

XX WPI; 2003-743764/70.

XX P-PSDB; ADC24126.

XX Novel polynucleotide sequence encoding Discosoma red fluorescent protein

XX variant having a reduced propensity to oligomerize, useful for detecting

XX transcriptional activity.

XX Disclosure; SEQ ID NO 9; 67pp; English.

CC The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (DsRed) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DsRed sequence (SI) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cell  
CC containing a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localization or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcriptions, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca<sup>2+</sup>, Zn<sup>2+</sup>, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is

CC useful for identifying a region or condition that regulates the activity  
 CC of an expression control sequence. This sequence represents a modified  
 CC Diacosoma wild-type red fluorescent protein coding region with humanised  
 CC codon usage.

XX  
 SQ Sequence 681 BP; 146 A; 223 C; 205 G; 107 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	7.26e-140	Length:	681
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	10	Gaps:	0

US-10-006-922a-12 (1-225) x ADC24134 (1-681)

```

QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20
DB 4 GTGGGCTCTTCAGAAAGCTCATCAGAGAGTTCAAGCTTCCAGTCCGATGAGAGGC 63
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40
DB 64 ACCGTGAACGGCCACGAGTTGAGATGAGGGGCGAGGGCGGCCCTTACGAGAGGC 123
QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60
DB 124 CACAAACCCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCCTGGGAGATC 183
QY 61 LeuSerProGlnPheGlnIyrgIySerIysValTyrrValIysHisProAlaAspIlePro 80
DB 184 CTGTCCCCCAAGTTCAGTACGGCTCCAGAGGTGACGTGAAGCAACCCCGCGACATCCCC 243
QY 81 AspTyrlsIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPhenGlu 100
DB 244 GACTACAAAGAGCTGCTCTCCCGAGGGCTTCAAGTGGGAGCGCGTGAATGAACTTCGAG 303
QY 101 AspGlyGlyValValThrValThrIAsnAspSerSerIeuGlnAspGlyCysePheIleTy 120
DB 304 GACGGCGCGCTGGGACCGTGAACCAAGACTCCCTCGACGAGCGCGCTGCTCATCTAC 363
QY 121 LysValIysPheIleGlyValIAsnPheProSerAspGlyProValMetGlnIysIysThr 140
DB 364 AAGGTGAAGTTCATCGCGCTGAACCTTCCCTCCGACGGCCCGCTTAATGCAGAAAGAAC 423
QY 141 MetGlyTrpGluIAsnSerThrGluArgLeuTyrrProArgAspGlyValLeuIysGlyGlu 160
DB 424 ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGAAGAGGCGAG 483
QY 161 IleHisIysAlaLeuIysLeuIysAspGlyGlyHisIyrlsLeuValGluPheIysSerIle 180
DB 484 ATCCACAAAGGCCCTGAAGCTGAAGAGCGGCGCCACTACCTGGTGAAGTTCAAGTTCATC 543
QY 181 TyrrMetAlaIysIysProValGlnLeuProGlyTyrrTyrrValIAspSerIysLeuAsp 200
DB 544 TACATGGCCAGAAAGCCCGTGCACCTGCGCGCTTAATCACTACGTGAGCTCCAGCTGAC 603
QY 201 IleThrSerHisAsnGluAspTyrrThrIleValGluGlnIyrgIuArgThrGluGlyArg 220
DB 604 ATGACCTCCCAACAGAGACTACACCATCTGAGAGGTACGAGCGCAGCGAGGGCGCG 663
QY 221 HisHisIeuPheIeu 225
DB 664 CACCACCTGTTCTTG 678

```

#### RESULT 24

ADL46205  
 ID ADL46205 standard; DNA; 681 BP.  
 AC ADL46205;  
 XX  
 DT 20-MAY-2004 (first entry)  
 .XX

DE Human codon optimised Diacosoma red fluorescent protein (Dered) DNA.  
 XX db; gene; red fluorescent protein; Dered; fluorescence; red wavelength;  
 KW oligomerization; tetramerization; immunoassay; hybridization assay.

XX  
 OS Diacosoma sp.

XX  
 PN WO2003086446-A1.

XX  
 PD 23-OCT-2003.

XX  
 PF 09-APR-2003; 2003WO-US010879.

XX  
 PR 10-APR-2002; 2002US-00121258.

XX  
 PR 29-JUL-2002; 2002US-00209208.

XX  
 PA (REGC ) UNIV CALIFORNIA.

XX  
 PI Telen RY, Campbell RE, Balrd GS;

XX  
 DR WPI; 2003-845265/78.

XX  
 PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
 PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
 PT in molecular biology, e.g. in immunoassays or in tracking protein  
 PT movement in cells.

XX  
 PS Disclosure; SEQ ID NO 3; 166pp; English.

XX  
 CC The invention relates to a polynucleotide sequence encoding a Diacosoma  
 CC red fluorescent protein (Dered) variant having a reduced propensity to  
 CC oligomerize. The protein variant comprises one or more amino acid  
 CC substitutions at the AB and/or AC interface(s) of the wild-type Dered  
 CC sequence, where the substitutions result in reduced propensity of the  
 CC Dered variant to form tetramers and where the variant displays detectable  
 CC fluorescence of at least one red wavelength. The composition and methods  
 CC are useful in producing red fluorescent proteins having reduced  
 CC propensity for oligomerization, especially tetramerization. The protein  
 CC may be used in molecular biology and in other scientific applications,  
 CC such as in immunoassays or hybridization assays, or in tracking the  
 CC movement of proteins in cells. This sequence corresponds to the Dered  
 CC coding sequence with codons optimised for human expression.

XX  
 SQ Sequence 681 BP; 146 A; 223 C; 205 G; 107 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	7.26e-140	Length:	681
Score:	1210.00 <td>Matches:</td> <td>224</td>	Matches:	224
Percent Similarity:	100.00% <td>Conservative:</td> <td>1</td>	Conservative:	1
Best Local Similarity:	99.56% <td>Mismatches:</td> <td>0</td>	Mismatches:	0
Query Match:	99.67% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	11	Gaps:	0

US-10-006-922a-12 (1-225) x ADL46205 (1-681)

```

QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20
DB 4 GTGGGCTCTTCAGAAAGCTCATCAGAGAGTTCAAGCTTCCAGTCCGATGAGAGGC 63
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40
DB 64 ACCGTGAACGGCCACGAGTTGAGATGAGGGGCGAGGGCGGCCCTTACGAGAGGC 123
QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60
DB 124 CACAAACCCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCCTGGGAGATC 183
QY 61 LeuSerProGlnPheGlnIyrgIySerIysValTyrrValIysHisProAlaAspIlePro 80
DB 184 CTGTCCCCCAAGTTCAGTACGGCTCCAGAGGTGACGTGAAGCAACCCCGCGACATCCCC 243
QY 81 AspTyrlsIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPhenGlu 100

```

Db 244 GACTACAAAGAGCTGTCCTTCCCGAGGCTTCAAGTGGAGCCGCTGATGAACTTCGAG 303  
 Qy 101 AspGlyValValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 Db 304 GACGGCGCGCTGTGACCGCTGACCCAGACTCTCCCTGACGAGACGGCTGCTCATCTAC 363  
 Qy 121 LysValIysPheIleGlyValIAsnPheProSerAspGlyProValMetGlnLysLeuThr 140  
 Db 364 AAGGTGAAGTTCACTCGCGCTGAACCTTCCCTCCACCGCGCCCTGTAAGTCAGAAAGAAC 423  
 Qy 141 MetGlyTyrProGlnIAsnSerThrGlnArgLeuTyrProArgAspGlyValIleuLysGlyIu 160  
 Db 424 ATGGGCTGGAGGCTCTCACCAGACGCTGTATCCCGCGAGCGCTGTGAAGGCGAG 483  
 Qy 161 ILeHisValIAsnLysLeuLysLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
 Db 484 ATCCACAAAGCCCTCGAAGCTGAAAGACGGCGCCACATCCTGTGTGAGATTCAATCCATC 543  
 Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValIAspSerLysLeuAsp 200  
 Db 544 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCTACTACTAGTGACTCCAAAGCTGGAC 603  
 Qy 201 ILeThrSerHisAsnGlnAspTyrThrIleValGluGlnTyrGlnArgThrGluGlyArg 220  
 Db 604 ATCACTCTCCCAACAAGAGACTACACATCGTGAAGCAGTACAGCGCACCGAGGCGCGC 663  
 Qy 221 HisHisLeuPheLeu 225  
 Db 664 CACCACTGTTCTCTG 678  
 Db  
 RESULT 25  
 ADL46225  
 ID ADL46225 standard; DNA; 681 BP.  
 AC ADL46225;  
 XX 20-MAY-2004 (first entry)  
 DT  
 XX Human codon optimised Discosoma red fluorescent protein (DRed) DNA.  
 DE  
 XX  
 KW de; gene; red fluorescent protein; DRed; fluorescence; red wavelength;  
 KW oligomerization; tetramerization; immunoassay; hybridization assay.  
 XX  
 OS Discosoma sp.  
 OS  
 PN WO2003086446-A1.  
 PD 23-OCT-2003.  
 PD  
 XX 09-APR-2003; 2003WO-US010879.  
 PF  
 XX 10-APR-2002; 2002US-00121258.  
 PR 29-JUL-2002; 2002US-00209208.  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Tsien RY, Campbell RE, Baird GS;  
 DR WPI; 2003-845265/78.  
 XX  
 XX New monomeric and dimeric Anthozoan fluorescent protein variants with  
 PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
 PT in molecular biology, e.g. in immunoassays or in tracking protein  
 PT movement in cells.  
 XX  
 XX Disclosure; SEQ ID NO 23; 166bp; English.  
 PS  
 XX The invention relates to a polynucleotide sequence encoding a Discosoma  
 CC red fluorescent protein (DRed) variant having a reduced propensity to  
 CC oligomerize. The protein variant comprises one or more amino acid  
 CC substitutions at the AB and/or AC interface(s) of the wild-type DRed  
 CC sequence, where the substitutions result in reduced propensity of the  
 CC dimered variant to form tetramers and where the variant displays detectable

CC fluorescence of at least one red wavelength. The composition and methods  
 CC are useful in producing red fluorescent proteins having reduced  
 CC propensity for oligomerization, especially tetramerization. The protein  
 CC may be used in molecular biology and in other scientific applications,  
 CC such as in immunoassays or hybridization assays, or in tracking the  
 CC movement of proteins in cells. This sequence corresponds to the DRed  
 CC coding sequence with codons optimised for human expression.  
 XX  
 SQ Sequence 681 BP; 146 A; 223 C; 205 G; 107 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 7,26e-140 Length: 681  
 Score: 1210.00 Matches: 224  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.56% Mismatches: 0  
 Query Match: 99.67% Indels: 0  
 DB: 11 Gaps: 0  
 US-10-006-922A-12 (1-225) x ADL46225 (1-681)  
 Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
 Db 4 GTGGCTCTCTCCAAAGAGCTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGAGGC 63  
 Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyValGluGlyArgProTyrGlyGly 40  
 Db 64 ACCGTGAACGCCCAAGTTCAGATCAGAGGCGAGGCGAGGCGCCCTTACGAGGC 123  
 Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
 Db 124 CACAAACCCGTGAAGCTGAAGGTGACCAAGGCGGCCCTCGCTTCGCTGGAGCATC 183  
 Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
 Db 184 CTGTCCCCCAGATTCCAGTACGCGCTCCAAAGGTGATGAAAGCACCCCGCGACATCCCC 243  
 Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysThrGluValMetAsnPheGlu 100  
 Db 244 GACTACAAAGAGCTGCTCTTCCCGAGGCTTCAAGTGGAGCGCTGATGAACTTCGAG 303  
 Qy 101 AspGlyValValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 Db 304 GACGGCGCGCTGTGACCGCTGACCCAGACTCTCCCTGACGAGACGGCTGCTCATCTAC 363  
 Qy 121 LysValIysPheIleGlyValIAsnPheProSerAspGlyProValMetGlnLysLeuThr 140  
 Db 364 AAGGTGAAGTTCACTCGCGCTGAACCTTCCCTCCACCGCGCCCTGTAAGTCAGAAAGAAC 423  
 Qy 141 MetGlyTyrProGlnIAsnSerThrGlnArgLeuTyrProArgAspGlyValIleuLysGlyIu 160  
 Db 424 ATGGGCTGGAGGCTCTCACCAGACGCTGTATCCCGCGAGCGCTGTGAAGGCGAG 483  
 Qy 161 ILeHisValIAsnLysLeuLysLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
 Db 484 ATCCACAAAGCCCTCGAAGCTGAAAGACGGCGCCACATCCTGTGTGAGATTCAATCCATC 543  
 Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValIAspSerLysLeuAsp 200  
 Db 544 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCTACTACTAGTGACTCCAAAGCTGGAC 603  
 Qy 201 ILeThrSerHisAsnGlnAspTyrThrIleValGluGlnTyrGlnArgThrGluGlyArg 220  
 Db 604 ATCACTCTCCCAACAAGAGACTACACATCGTGAAGCAGTACAGCGCACCGAGGCGCGC 663  
 Qy 221 HisHisLeuPheLeu 225  
 Db 664 CACCACTGTTCTCTG 678  
 Db  
 RESULT 26  
 ABZ22476  
 ID ABZ22476 standard; DNA; 723 BP.  
 AC ABZ22476;

```

XX 25-MAR-2003 (first entry)
XX Mammalian codon optimised Discosoma red fluorescent protein DNA.
XX Mammalian codon optimised Discosoma red fluorescent protein; Discosoma;
XX red fluorescent protein; directed evolution; gene; ds.
XX
XX Discosoma sp.
XX Mammalia.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 1..723
XX /*tag= a
XX /product= "Mammalian codon optimised Discosoma red
XX fluorescent protein"
XX /transl_except= (pos:616..618,aa:His)
XX
XX MO200294992-A2.
XX
XX 28-NOV-2002.
XX
XX 20-MAY-2002; 2002MO-US015968.
XX
XX 18-MAY-2001; 2001US-0291871P.
XX (RIGB-) RIGEL PHARM INC.
XX
XX Peelle B;
XX
XX WPI; 2003-120798/11.
XX P-PSDB; ABP56678.
XX
XX New Discosoma red fluorescent protein, useful for functional screens as a
XX reporter for gene transcription, for target characterization and
XX localization of fusion proteins, or for scaffolds for protein and peptide
XX libraries.
XX
XX Example 1; Fig 1; 22pp; English.
XX
XX The present invention describes an isolated Discosoma red fluorescent
XX protein (1) comprising a 241 residue amino acid sequence (see ABP56678,
XX S1), with one or more point mutations at amino acid position N24, F125,
XX K164, or M183. Also described: (1) a fusion protein comprising (1); (2)
XX an isolated nucleic acid encoding (1); (3) a vector comprising the
XX nucleic acid of (2); (4) a host cell comprising the vector; (5) a
XX retroviral cDNA expression library comprising the nucleic acid of (2), or
XX encoding (1); and (6) methods of making a fluorescent variant. The
XX Discosoma red fluorescent proteins are useful for functional screens as a
XX reporter for gene transcription (e.g. as a fusion protein), for target
XX characterisation and localisation of fusion proteins, or for scaffolds
XX for protein and peptide libraries. The fluorescent proteins can also be
XX used as selectable markers or reporter molecules for a variety of
XX bioassays, including methods that use fluorescence activated cell sorting
XX (FACS) as a selection mechanism. The method of directed protein evolution
XX is useful for obtaining improved variants of red fluorescent protein. The
XX variants of Discosoma red fluorescent protein have greatly improved
XX brightness, expression, and/or folding kinetics as compared to wild type
XX or a codon optimised variant. The present sequence encodes a mammalian
XX codon optimised Discosoma red fluorescent protein, which is used in an
XX example from the present invention
XX
XX Sequence 723 BP; 161 A; 233 C; 221 G; 108 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 7.9e-140 Length: 723
XX Score: 1210.00 Matches: 224
XX Percent Similarity: 100.00% Conservative: 1
XX Best Local Similarity: 99.56% Mismatches: 0
XX Query Match: 99.67% Indels: 0
XX DB: 8 Gaps: 0

```

```

US-10-006-922a-12 (1-225) x AB222476 (1-723)
QY 1 MetArgSerSerIyAsnValIlelygSluphMetArgPheIyValArgMetGluY 20
DB 4 GTGCGCTCTCCCAAGACGTCATCAAGAGGAGTTTCATCGCTTCAAGGTCGATGAGGCG 63
QY 21 ThrValAsnGlyYHIsGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40
DB 64 ACCGTGAACGCGCAACAGATTCCAGATCGAGGGGAGGCGGCGCCCTTCAGAGGCG 123
QY 41 HisAsnThrValIyAsnLeuIyValIThrlysgIyGlyProIeupProPheIatPAPPIle 60
DB 124 CACAACACCGTAGAGCTGAAGGTGACCAAGGGGGGCCCTTCGCTTCGCTGGGACATC 183
QY 61 LeuSerProGluPheGluIlyTyrGlySerIyValIyValIleHisProIaAspIlePro 80
DB 184 CTGTCCCCCAAGTTCACAGTCAAGGCTCCAGAGTGTACGTGAAGCAACCCCGCATCCCC 243
QY 81 AspTyrIleIyAsnLeuSerPheProGluGlyPheIyETripGluArgValMetAsnPhGlu 100
DB 244 GACTTCAAGAGAGCTGTCTTCCCGAGGCTTCAGTGGAGGCGGTGATGAATTCGAG 303
QY 101 AspGlyGlyValValIThrValIThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120
DB 304 GACGGCGGCGGTGACCGGTGACCGGACCAAGACTCTCCCTCGACGAGACGCGTTCATCTAC 363
QY 121 LysValIyPheIleIyValIAsnPheProSerAspGlyProValMetGluIyIyThr 140
DB 364 AAGGTGAAGTTCATCGCGGTGAACTTCCCTCGAGCGGCCCTGAATGCAAGAAAGACC 423
QY 141 MetGlyTyrGluIyAsnSerThrGluArgLeuTyrProArgAspGlyValIleuIysgIyGlu 160
DB 424 ATGGCGTGGAGGCGCTCCACCGAGCGCCCTGTACCCCGCAGACGCGGTGAGGCGGAG 483
QY 161 IleHisIyAsnIleuIyAsnLeuIyAspGlyGlyYHIsTyrIleuValGluPheIySerIle 180
DB 484 ATCCACAAGGCGCTGAAGCTGAAGAGCGGCGGCACCTACCTGTGAGTTCAAGAGTATC 543
QY 181 TyrMetAlaIyAsnProValGlnIeupProGlyTyrTyrTyrValAspSerIyIyAsp 200
DB 544 TACATGCGCAAGAGCGCGGTGAGCTGCCGCTACTACTACGTGACCTCAAGCTGAGC 603
QY 201 IleThrSerHisAsnGluAspTyrThrIleValIleGluGluTyrGluArgThrGluIyArg 220
DB 604 ATCACTCCCAACAGAGACTACACCATCTGTGAGACAGTACAGAGCGCACGAGGCGCG 663
QY 221 HisHisIeupHeuIeu 225
DB 664 CACCACCTGTTCTCG 678
RESULT 27
ADL18131
ID ADL18131 standard; DNA; 1638 BP.
XX
XX ADL18131;
XX
XX 06-MAY-2004 (first entry)
XX
XX RFP:PS(N1a protease):AtOBP7:GFP fusion protein DNA SEQ ID NO:51.
XX
XX chimeric protein; signal protein; trafficking signal targeting;
XX proteolytic cleavage site; protease; protease inhibitor; gene; ds.
XX
XX Arabidopsis thaliana.
XX
XX Tobacco vein motting virus.
XX
XX Synthetic.
XX
XX MO2003014381-A1.
XX
XX 20-FEB-2003.
XX
XX 08-AUG-2002; 2002MO-KR001515.
XX

```



```
PR 10-AUG-2001; 2001KR-00048123.
XX
XX (AHRA-) AHRAM BIOSYSTEMS INC.
XX
XX Hwang I, Kim DH, Lee YJ;
PI WPI; 2003-256596/25.
XX
XX P-PSDB; ADL18132.
XX
PT New chimeric protein, useful for detecting protease inhibitors inside the
XX cell or tissue.
XX
XX Example 2; SEQ ID NO 51; 214pp; English.
XX
XX The present invention describes a chimeric protein comprising at least
XX one signal protein that has a trafficking signal targeting to a
XX subcellular organelle and at least one proteolytic cleavage site for a
XX protease. The chimeric protein is constructed, so that: (a) the
XX trafficking signals of all the signal proteins are inactivated by linking
XX the proteolytic site or a signal masking protein through the proteolytic
XX site to the N- or C- terminus of the signal proteins, and so the chimeric
XX protein is present in cytosol; (b) the trafficking signal of at least one
XX signal protein is activated when the proteolytic cleavage site is cleaved
XX by the protease, and as a result at least one fragment protein that
XX includes the activated signal protein is a transported to a subcellular
XX organelle; and (c) the chimeric protein is a transported to a subcellular
XX fluorescent label signal in the cell is altered depending on the cleavage
XX by the protease. Also described: (1) a recombinant gene comprising a
XX nuclear acid sequence encoding the chimeric protein which is constructed
XX to express the chimeric protein in a cell; (2) a cell transformed with
XX the recombinant gene or vector; (3) analysing the activity of a protease
XX in vivo; (4) screening protease inhibitors in vivo; (5) a system for
XX detecting a protease inside a cell; (6) a nucleic acid comprising the
XX sequence encoding the chimeric protein for detecting protease activity in
XX a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting
XX a protease inside a cell comprising the chimeric protein or the vector;
XX (9) detecting a protease inside a cell or tissue; and (10) detecting a
XX protease inhibitor in vivo. The chimeric protein is useful for detecting
XX protease inhibitors inside the cell or tissue. The present sequence
XX encodes a fusion protein, which is used in the exemplification of the
XX present invention.
XX
SQ Sequence 1638 BP; 445 A; 450 C; 419 G; 324 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.5e-139 Length: 1638
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservatave: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x ADL18131 (1-1638)
QY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20
DB 4 GTGGCTCTCTCCAAAGACGTCTCAAGAGTTCAAGCGTTCAAGSTGGCATGAGGGCC 63
QY 21 ThrValaengLyHISgluPheGluIleGluGlyGluGlyValArgProTyrgluGly 40
DB 64 ACCGTGAACGGCCAGAGTTCCAGATCCAGAGCGCGGAGGGCCGCCCTTCCACAGAGGCC 123
QY 41 HisAsnThrValLyLeuLyValPheTrlyGlyValProLeuProPheLatrapapile 60
DB 124 CACCAACACCGTAAAGGTGACCAAGAGGGCGGGCCCTCTCCCTCGGAGCATTC 183
QY 61 LeuSerProGlnPheGlnTyrgLySerLyValTyValLyHISProLAAspIlePro 80
DB 184 CTGTCTCCCACTTCCAGTACGGCTCCAGAGTGTACGTAAACACCCCGCCGACATCCCC 243
QY 81 AspTyTyValLyLeuSerPheProGluGlyPheLyValTyrgLyValArgValMetAsnPhGlu 100
```

```
DB 244 GACTACAGAAGAGCTGTCTTCCCGAGGGCTTCAGTGGAGCGCGTATGAATTCGAG 303
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySpheIleTy 120
DB 304 GACGCGCGCTGTGTGACCGTGAACCCAGACTCTCTCCGTGAGAGACGCGCTCTCATCTAC 363
QY 121 LyValLySpheIleGlyValaAsnPhPheProSerAspGlyProValMetGlnLySpheThr 140
DB 364 AAGGTGAAGTTCAATCGCGTGAACCTTCCCTCCGACGGCCCGCTTATGCAAGAAAGACC 423
QY 141 MetGlyTyrgLyAspSerThrGluArgLeuTyProArgAspGlyValaLeuLyGlyGlu 160
DB 424 ATGGCGTGGAGAGGCTCCACAGAGCGCGCTGATACCCCGCAGCGCGTGAAGGGCGAG 483
QY 161 IleHisLyAlaLeuLyLeuLyAspGlyGlyHISTyLeuValGluPheLySerIle 180
DB 484 ATCCACAAAGCCCTGGAAGCTGAAGAGAGCGGGCCACTACTGAGTTCATCAATCATC 543
QY 181 TyrMetAlaLyLyAspProValGlnLeuProGlyTyTyTyTyValaAspSerLyLeuAsp 200
DB 544 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCTACTACTACGTGAGCTCAAGCTGAGC 603
QY 201 IleThrSerHisAsnGluAspTyTyThrIleValGluGlnTyrgLyValArg 220
DB 604 ATCACTCTCCACAAACAGAGACTACACATGTGTGAGAGTACGAGCCGACCGAGCGCG 663
QY 221 HisHisLeuPheLeu 225
DB 664 CACCACCTGTCTCTG 678

RESULT 28
ADL18155
ID ADL18155 standard; DNA; 1647 BP.
XX
XX AC ADL18155;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE RFP:PS (HIV-1 protease); AtOEFP7:GFP fusion protein DNA SEQ ID NO:75.
XX
XX KW chimeric protein; signal protein; trafficking signal targeting;
XX proteolytic cleavage site; protease; protease inhibitor; gene; ds.
XX
XX OS Arabidopsis thaliana.
XX
XX OS Human immunodeficiency virus 1.
XX
XX OS Synthetic.
XX
XX PN WO2003014381-A1.
XX
XX PD 20-FEB-2003.
XX
XX PF 08-AUG-2002; 2002WO-KR001515.
XX
XX PR 10-AUG-2001; 2001KR-00048123.
XX
XX PA (AHRA-) AHRAM BIOSYSTEMS INC.
XX
XX PI Hwang I, Kim DH, Lee YJ;
XX
XX WPI; 2003-256596/25.
XX
XX P-PSDB; ADL18156.
XX
XX PT New chimeric protein, useful for detecting protease inhibitors inside the
XX cell or tissue.
XX
XX Example 3; SEQ ID NO 75; 214pp; English.
XX
XX The present invention describes a chimeric protein comprising at least
XX one signal protein that has a trafficking signal targeting to a
XX subcellular organelle and at least one proteolytic cleavage site for a
XX protease. The chimeric protein is constructed, so that: (a) the
XX trafficking signals of all the signal proteins are inactivated by linking
XX the proteolytic site or a signal masking protein through the proteolytic
```

CC site to the N-or C-terminus of the signal protease, and so the chimeric  
CC protein is present in cytosol; (b) the trafficking signal of at least one of  
CC signal protein is activated when the proteolytic cleavage site is cleaved  
CC by the protease, and as a result at least one fragment protein that  
CC includes the activated signal protein is transported to a subcellular  
CC organelle; and (c) the chimeric protein is labelled with at least one  
CC fluorescent protein and the position and intensity distribution of the  
CC fluorescent label signal in the cell is altered depending on the cleavage  
CC by the protease. Also described: (1) a recombinant gene comprising a  
CC nucleic acid sequence encoding the chimeric protein which is constructed  
CC to express the chimeric protein in a cell; (2) a cell transformed with  
CC the recombinant gene or vector; (3) analysing the activity of a protease  
CC in vivo; (4) screening protease inhibitors in vivo; (5) a system for  
CC detecting a protease inside a cell; (6) a nucleic acid comprising the  
CC sequence encoding the chimeric protein for detecting protease activity in  
CC a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting  
CC a protease inside a cell comprising the chimeric protein or the vector;  
CC (9) detecting a protease inside a cell or tissue; and (10) detecting a  
CC protease inhibitor in vivo. The chimeric protein is useful for detecting  
CC protease inhibitors inside the cell or tissue. The present sequence  
CC encodes a fusion protein, which is used in the exemplification of the  
CC present invention.

SQ Sequence 1647 BP; 450 A; 448 C; 421 G; 328 T; 0 U; 0 Other;

**Alignment Scores:**

Pred. No.:	2.52e-139	Length:	1647
Score:	1210.00	Matches:	224

Score:	1210.00	Matches:	224
--------	---------	----------	-----

Percent Similarity: 100.00%      Conservative: 1

Best Local Similarity: 99.56% Mismatches: 0

Query Match:	99.67%	Indels:	0
--------------	--------	---------	---

```
DB: 10 Gaps: 0
```

US-10-006-922A-12 (1-225) X ADL18155 (1-1647)

QY	1	MetAAGSesSerLyAaAntValIleLysGIuphMetAaRgPheLysValArgMetGIuGI	20
Db	4	GTGGCCCTCTCCAAAGAACTCATCAAGAGATTATGGCTTCAAAGTGCCATGAGAGGC	63
QY	21	ThrValaenGIvHIsGIuphneGIuIleGIuGIyGIuGIyGIuGIyAaProTyRGIuGI	40
Db	64	ACCGGAACGGCCAGAGTTCCGAGATCGAGGGGAGGGGCGAGGGCCGCTTCGAGAGGC	123
QY	41	HisAntThrValaLysLeuLysValThrLysGIyGIyProLeuProPheAlaTPaerIle	60
Db	124	CACAAACACCGTGAAGGTGAAGGGACCAAGGGGGGGCCCTGGCTTCGCTGGAGATC	183
QY	61	LeuSerProGIuPhneGIuTyRGIySerLysValTyRValLysIAsProIAsaPIlePro	80
Db	184	CTGTCCCCCAAGTTCAGTACGGCTCCAAAGTATAGTAAAGCACCCCGCCAGATCCCC	243
QY	81	AerTyTLysLysLeuSerPheProGIuGIyPheLysTPGIuAaRgValMetAaPheGIu	100
Db	244	GACCTCAAGAACTGCTCTTCCCGAGGGCTTCAAATGGGAGCGGTGAATCTTCGAG	303
QY	101	AerGIyGIyValValThrValThrGIuAaPseSerLeuGIuAaPbGIyCysPheIleTyR	120
Db	304	GACGGCGGCTGATGACCGGTGAACCAAGACTCTCCCTGAGGAAGCGCTTCATCTAC	363
QY	121	LysValLysPheIleGIyValAaPheProSerAaPbGIyProValMetGIuLysThr	140
Db	364	AAGGTGAAGTTCAATCGGGCTGAACCTCCCTCCGAGGGCCCGTAAATGCAAGAAAGACC	423
QY	141	MetGIyTPGIuAaSerThrGIuAaGLeuTyRProAaAerGIyValLeuLysGIyGIu	160
Db	424	ATGGCTGTGGAGGCTCCACCGAGGCCCTGTACCCCGCAGCGCGTGAAGGGGGAG	483
QY	161	IleHisLysValaLeuLysLysAaPbGIyGIyHisTyRLeuValGIuPheLysSerIle	180
Db	484	ATCCACAAAGCCCTGAAGCTGAAGAGACGGCGGCGCATCACTGTGTGAAGTTCAAGTCAATC	543
QY	181	TyRMetAlaLysLysProValGIuLeuProGIyTyTyTyRValAaPseSerLysLeuAaP	200

Db	544	TACATGCGCAAGAAAGCCCGCTGACGTGCCCGGCTACTACTACTGATGATCCAAAGCTGGAC	603		
Qy	201	lIeThSeRiHsAsNgUaSPYrThrIlleValGluInTYrGluarGThrgluGIyarg	220		
Db	604	ATCACCCTCCCAACAGAGACTACACCATCGGAGACAGTACGAGGCGACCGAGGCGCG	663		
Qy	221	HiSHiSlauPhelu	225		
Db	664	CACCACCTGTCCTG	678		
RESULT 29					
ID	AD575466	standard; DNA; 2721 BP.			
XX	AD575466;				
XX	02-DEC-2004	(first entry)			
XX					
De		Fibrohexamerin promoter; exon 1 and intron 1 synthetic sequence SEQ ID 1.			
XX					
Kw		Spidroin; fibroin; silkworm; Ds-Red; Fibrohexamerin; gene; promoter; ds.			
XX					
OS		Synthetic.			
FH	Key	Location/Qualifiers			
FT	promoter	1..1452			
FT		/*tag= a			
FT		/label= Fibrohexamerin promoter			
FT	misc_feature	1150..2026			
FT		/*tag= b			
FT		/note= "This region is specifically claimed in Claim 2"			
FT	CAAT_signal	1362..1365			
FT		/*tag= c			
FT	misc_signal	1379..1390			
FT		/*tag= d			
FT		/note= "Binding site for transcription factors of the			
FT		SGF1 forkhead type"			
FT	TATA_signal	1420..1423			
FT		/*tag= e			
FT	exon	1452..1525			
FT		/*tag= f			
FT		/number= 1			
FT		/note= "Fibrohexamerin exon 1, encodes the signal			
FT		peptide"			
FT	intron	1526..2026			
FT		/*tag= g			
FT		/number= 1			
FT		/note= "Fibrohexamerin intron 1"			
FT		2041..2721			
FT	CDS	/*tag= h			
FT		/product= "Reporter protein Ds-Red"			
PN	FR2852325-A1.				
XX					
PD	17-SEP-2004.				
XX					
PF	13-MAR-2003; 2003FR-00003137.				
XX					
PR	13-MAR-2003; 2003FR-00003137.				
XX					
PA	(CNRS ) CNRS CENT NAT RECH SCI.				
PA	(UYLY-) UNIV LYON 1 BERNARD CLAUDE.				
PA	(INRG ) INRA INST NAT RECH AGRONOMIQUE.				
XX					
P1	Chavancy G, Couble P, Durand B, Grenier A, Horard B, Julien B;				
XX					
XX	Mauchamp B, Nony P, Prudhomme JC, Royer C;				
XX					
DR	WPI; 2004-670997/66.				
XX					
PT	New nucleic acid that directs expression of protein in silk glands of				
PT	silkworm, useful for preparation of pharmaceutical proteins and for				
XX	modifying textile properties of silk.				

PS Claim 1; SEQ ID NO 1; 50pp; French.

XX The present invention relates to a novel nucleic acid (I) that directs  
CC expression of a protein of interest (II) specifically in the cells of the  
CC posterior silk-producing glands (A) of Bombyx mori. (I) comprises, in the  
CC 5' to 3' direction: a regulatory region (RI) comprising signals for  
CC regulating expression of a polynucleotide of interest (III) in the  
CC specified cells and, under control of RI, a region that encodes a  
CC modified signal peptide (msp). (I) is at least 90% identical with the  
CC 1150-2026 nucleotide (nt) region of AD575466, in which the codon  
CC containing nt 1486-1488 encodes Ala, Ile or Leu. (I) is optionally  
CC modified by insertion of 1-4 copies of the sequence AD575467 (nucleotides  
CC 1379-1390 of AD575466), which is the binding site for transcription  
CC factors of the SGF1 forkhead type, functioning as activator of  
CC transcription and regulator of tissue-specific expression. (III) encodes  
CC a polypeptide, specifically spideroin (of spiders) or fibroin (from  
CC organisms of the genus Galleria). (I), also related expression cassettes  
CC and vectors, are used to prepare transgenic Bombyx that secrete (II)  
CC incorporated into silk threads, e.g. where (II) is of biomedical interest  
CC (e.g. hormones, antigens, enzymes, growth factors and receptors) or where  
CC the threads are intended for use as textiles and (II) modifies resistance  
CC to ductility or elasticity.

XX Sequence 2721 BP; 748 A; 631 C; 606 G; 736 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	5.11e-139	Length:	2721
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	13	Gaps:	0

US-10-006-922a-12 (1-225) X AD575466 (1-2721)

QY 1 MecarGSeSerSerLeuAenValIleLyGluPheNeAArgPheLyValArgMetGluGly 20  
DB 2044 GTGGCGCTCTCCAGAGAACGTCAAGAGATTCAAGCCCTTCAGAGTCGATGAGGCG 2103  
QY 21 ThrValaAngLyHleGluPheGluIleGluGlyGluGlyGluGlyAArgProGlyGluGly 40  
DB 2104 ACCGTGAACGGCCAGAGTTCAAGATCAAGGCGGAGGCGGAGGCGGCGGCTTACAGAGGC 2163  
QY 41 HisAenThrValLyLeuLyValThrLyGlyGlyProLeuProPheAlaITraAspIle 60  
DB 2164 CACCAACACCGTGAAGCTGACGATGACCAAGGCGGCGGCGGCTTCCGCTGGACATC 2223  
QY 61 LeuSerProGlnPheGlnTrGlySerLyValTrValLyHleProAlaAspIlePro 80  
DB 2224 CTGTCCCGCCAGTTCCAGTACGGCTCAAGGTGTACGTGAACACCCGCGCCATCCCC 2283  
QY 81 AspTrLyLyLeuSerPheProGluGlyPheLyValArgValMetAsnPheGlu 100  
DB 2284 GACTACAAAGAGCTCTCTCCCGAGGCTTCAAGTGAAGCGGCTGATGAATTCGAG 2343  
QY 101 AspGlyLyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyAspHeIleTr 120  
DB 2344 GACGCGCGGCTGTACCTGACCCAGACATCTCTCCCTGCAAGAGCGGCTTCAATCAC 2403  
QY 121 LySValLyPheHeIleGlyValaAnPheProSerAspGlyProValMetGlnLySValThr 140  
DB 2404 AAGGTGAAGTTCAATCGCGCTGAATCTCCCTCCGACGCGGCGGCTATGCAAGAAAGACC 2463  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTrProArgAspGlyValLeuLySValGlu 160  
DB 2464 ATGGGCTGGAGAGCCCTCCACGAGCCCTGTACCCCGGAGGCGGTGCTGAAGGCGAG 2523  
QY 161 IleHleLyAlaLeuLyLeuLyAspGlyGlyHleTrLeuValGluPheLySerIle 180  
DB 2524 ATCCACAAAGCCCTGAAGCTGAAGAGCGGCGGCACTACCTGTGAGATTCAAGTCAATC 2583  
QY 181 TrMetAlaLyLyLeuProValGlnLeuProGlyTrTrTrValaAspSerLySValuAsp 200

DB 2584 TACATGCGCAAGAACCCCGTGCAGCTGCCCGCTACTACTACGTGACATCCAACTGAGAC 2643  
QY 201 IleThrSerHisAenGluAenValIleLyGluPheNeAArgPheLyValArgMetGluGly 220  
DB 2644 ATCACTCCCAACAGAGGACTACACCATGTGAGAGATGACAGCGCACCGAGGCGGC 2703  
QY 221 HisHlePhePheLeu 225  
DB 2704 CACCACTGTCTCTG 2718  
RESULT 30  
AD575468  
ID AD575468 standard; DNA; 2772 BP.  
XX  
AC AD575468;  
XX  
XX 02-DEC-2004 (first entry)  
DB Fibrohexamerin promoter, exon 1 and intron 1 synthetic sequence SEQ ID 3.  
XX  
XX Spideroin; fibroin; silkworm; Ds-Red; Fibrohexamerin; gene; promoter; ds.  
XX  
OS Synthetic.  
XX  
FH Key  
FH Location/Qualifiers  
FH 1..1503  
FT promoter  
FT /\*tag= a  
FT /label= Fibrohexamerin promoter  
FT 1362..1365  
FT CANT\_signal  
FT /\*tag= b  
FT 1371..1440  
FT misc\_signal  
FT /\*tag= c  
FT /note= "Three copies of the binding site for  
FT /note= transcription factors of the SGF1 forkhead type"  
FT 1471..1474  
FT /\*tag= d  
FT 1503..1576  
FT exon  
FT /\*tag= e  
FT /number= 1  
FT /note= "Fibrohexamerin exon 1, encodes the signal  
FT peptide"  
FT 1577..2077  
FT /\*tag= f  
FT /number= 1  
FT /note= "Fibrohexamerin intron 1"  
FT 2092..2772  
FT CDS  
FT /\*tag= g  
FT /product= "Reporter protein Ds-Red"  
XX  
FN FR2852325-A1.  
XX  
PD 17-SEP-2004.  
XX  
PF 13-MAR-2003; 2003FR-00003137.  
XX  
PR 13-MAR-2003; 2003FR-00003137.  
XX  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
PA (UVELY-) UNIV LYON 1 BERNARD CLAUDE.  
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX  
XX Chavancy G, Couble P, Durand B, Grenier A, Horard B, Julien B;  
PI Mauchamp B, Nony P, Prudhomme JC, Royer C;  
XX  
DR WPI: 2004-670997/66.  
XX  
XX New nucleic acid that directs expression of protein in silk glands of  
PT silkworm, useful for preparation of pharmaceutical proteins and for  
PT modifying textile properties of silk.  
XX  
PS Example 3; SEQ ID NO 3; 50pp; French.  
XX The present invention relates to a novel nucleic acid (I) that directs

CC expression of a protein of interest (II) specifically in the cells of the  
CC posterior silk-producing glands (A) of Bombyx mori. (I) comprises, in the  
CC 5' to 3' direction: a regulatory region (RI) comprising signals for  
CC regulating expression of a polynucleotide of interest (III) in the  
CC specified cells and, under control of RI, a region that encodes a  
CC modified signal peptide (mSP). (I) is at least 90% identical with the  
CC 1150-2026 nucleotide (nt) region of AD575466, in which the codon  
CC containing nt 1486-1488 encodes Ala, Ile or Leu. (I) is optionally  
CC modified by insertion of 1-4 copies of the sequence AD575467 (nucleotides  
CC 1379-1390 of AD575466), which is the binding site for transcription  
CC factors of the SGF1 forkhead type, functioning as activator of  
CC transcription, specifically spiderin (of spiders) or fibroin (from  
CC a polypeptide, specifically spiderin (of spiders) or fibroin (from  
CC organisms of the genus Galleria). (I), also related expression cassettes  
CC and vectors, are used to prepare transgenic Bombyx that secrete (II)  
CC incorporated into silk threads, e.g. where (II) is of biomedical interest  
CC (e.g. hormones, antigens, enzymes, growth factors and receptors) or where  
CC the threads are intended for use as textiles and (II) modifies resistance  
CC to ductility or elasticity. The present sequence is a modified version of  
CC AD575466 which comprises three copies of the SGF1 forkhead type  
CC transcription factor binding site.

XX SQ Sequence 2772 BP; 760 A; 640 C; 609 G; 763 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 5.24e-139 Length: 2772  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 13 Gaps: 0

US-10-006-922a-12 (1-225) x AD575468 (1-2772)

QY 1 MetArgSerSerLybAnVAlleLyGluPheMeArGpHeLybValArgMeGluGly 20  
DB 2095 GTGGGCTCTCCCAAGAACGATCATCAAGAGTTTCAAGGCTTCAAGGTCGCGATGGAAGGC 2154  
QY 21 ThrValAnGlyHtIsGluPheGluIleGluGlyGluGlyArgProTyGluGly 40  
DB 2155 ACCGTGAACGGCCACGAGTTGAGATGAGGCGACGGCGCCCTTACGAGGCGC 2214  
QY 41 HisAnThrValLybLeuLybValThrLybGlyPheProLeuPheAlaTrpAspIle 60  
DB 2215 CACAAACCCGGAAGCTGAAGGTACCAAGGCGCGCCCTTCGCTGGAGAACATTC 2274  
QY 61 LeuSerProGlnPheGlnTyGlySerLybValTyValLybHisProAlaAspIlePro 80  
DB 2275 CTGTCCCCCAGTTCCAGTACGGCTCCAGGTGTACGTGAAGCAACCCCGCGACATCCCC 2334  
QY 81 AspTyLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAnPheGlu 100  
DB 2335 GACTACAAAGAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCCCGATGAACTTCGAG 2394  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTy 120  
DB 2395 GACGGCGCGGTGTGACCGTACCCAGACCTCTCCCTCGAGGACGGGTGTTATCTAC 2454  
QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybSerThr 140  
DB 2455 AAGGTGAAGTTCACTGGGTGAATCTTCCCTCCGACGGCCCCCGTAAAGCAAGAACCC 2514  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyTrpArgAspGlyValLeuLybGlyGlu 160  
DB 2515 ATGGGCTGTGGAGGCTCCACGACGCGCTGTACCCCGCGACGCGCTGTGAAGGCGCAG 2574  
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyTrpLeuValGluPheLybSerIle 180  
DB 2575 ATCCACAAGGCCCTGAAGCTGAAGTGAAGGCGCGCACTACCTGTGTGAAGTTCAATCATTC 2634  
QY 181 TyMetAlaLybLybProValGlnLeuPProGlyTyTrpTyTrpValAspSerLybLeuAsp 200  
DB 2635 TACATGGCCAAAGAACCCGTCGACGTCGCGCTTACTACTACGTGACTCCAAAGCTGGAC 2694

QY 201 IleThrSerHisAnGluAspTyTrpThrIleValGluGlnTyArgGluArg 220  
DB 2695 ATACCTCTCCCAAGAACGATCATCAAGAGTTTCAAGGCTTCAAGGTCGCGATGGAAGGC 2754  
QY 221 HisHisLeuPheLeu 225  
DB 2755 CACCACTGTCTCTG 2769

RESULT 31

AA147954 standard; DNA; 4692 BP.

AC AA147954;

DT 26-SEP-2002 (first entry)

DE Modified yeast strain related vector pDRRedd1-N1.

KM Yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette;

KM modified yeast strain; environmental pollution; vector; ds.

OS Synthetic.

PN DE10061872-A1.

PD 20-JUN-2002.

PF 12-DEC-2000; 2000DE-01061872.

PR 12-DEC-2000; 2000DE-01061872.

XX (LICR/) LICHTENBERG-FRATTE H.

PI Lichtenberg-Fratte H;

DR WPI; 2002-539633/58.

PT Modified yeast strain, useful for detecting toxic compounds in

PT environment, contains integrated cassettes responsive to genotoxic and

XX Example 1; Page 24-25; 34pp; German.

CC The present invention relates to a modified yeast strain that contains,

CC integrated stably and functionally in its genome, a genotoxicity cassette

CC and a cytotoxicity cassette, each comprising a promoter and reporter

CC gene, both of which are different in the two cassettes. The modified

CC yeast strain is used to detect environmental pollution, especially

CC genotoxic and/or cytotoxic substances in complex environmental

CC contaminants, especially organic compounds, but also (non-)ionising

CC radiation and chemical carcinogens. Particular applications are in

CC monitoring (waste) water (e.g. as an early warning system), medical

CC toxicology screening and for industrial process control. The present

CC sequence is a vector suitable for use in the present invention

XX Sequence 4692 BP; 1111 A; 1268 C; 1247 G; 1066 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.1e-138 Length: 4692  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AA147954 (1-4692)

QY 1 MetArgSerSerLybAnVAlleLyGluPheMeArGpHeLybValArgMeGluGly 20  
DB 682 GTGGGCTCTCCCAAGAACGATCATCAAGAGTTTCAAGGCTTCAAGGTCGCGATGGAAGGC 741  
QY 21 ThrValAnGlyHtIsGluPheGluIleGluGlyGluGlyArgProTyGluGly 40

Db		742	ACCGTGAAGCCGCAAGATTCGAATCAGAAGGCAGGGCCGCCCTTAAGAGGC	801		
Oy		41	HtAaAnthrValLysLeuLysValThrLysGIYgLYProLeuPProPhaIatrpAspIle	60		
Db		802	CACAACACCGTGAAAGCTGAAGGTACCAAGGCGCGCCCTCGCCCTTGCGTTGGAATC	861		
Oy		61	LeuSerProGlnInpHeIntYrGLYSerLysValTYrVallYshSPRoAlaaEPilEpro	80		
Db		862	CTGTCGCCCGCATTCAGTAGCGGCTCCCAAGGTGTACGTAAAGCACCCCGCCGATCCCC	921		
Oy		81	AapTYrLYrLyLysLeuSerPheProGluGlyPheLYSTrPGlUaGVALMetAanPheGlu	100		
Db		922	GACTACAAAGAAAGCTGTCTCTCCCGAAGGCTTCAAATGGAGCGCGGTGAATGACTTCGAG	981		
Oy		101	AapGIYglYValValThrValThrGlnaSPserSerLeuGlnaAPGIYCYPheIIeTYr	120		
Db		982	GACGGCGGCGGTGGTACCGGTACCAGAACCTTCCTCCGACAGAACGGGTGTTTATCTAC	1041		
Oy		121	LysValLYrPheIIeGlyValaAanPheProSerAapGIYProValMetGlnLYrLysThr	140		
Db		1042	AAGGTGAAGTTCATCGCGGTGAATTCCCCTCCGACGGCCCGTAAATGCAGAAAGACCC	1101		
Oy		141	MetGIYTrPGlUalaserThrGluArgLeuTYrProaAgAapGIYValLeuLYgLIYGLU	160		
Db		1102	ATGGCGCTGGAGGCGCTTCCACCGAGGCGCTGTACCCCGACGCGCGTGTAAAGGCGAG	1161		
Oy		161	ILeHtSLySValaleuLYrLeuLYrAapGIYgLIYHtLYrLeuValGluPheLYrSerIle	180		
Db		1162	ATCCAACAAGGCCCTGAAGCTGAAGACCGCGGCCACTACCTGGTGGAGTTCAAGTTCATC	1221		
Oy		181	TyrMetalalySLySProValGlnLeuProGIYTYrTYrTYrValaASPserLYrLeuaSP	200		
Db		1222	TACATGGCCAAAGAACCCGTGAGCTGCGCGGCTACTACTAGTGAATCCAAAGCTGGAC	1281		
Oy		201	IleThrSerHtaenGluaspTYrThrIleValGluGlnTYrGLUAgrThrGluGIYArg	220		
Db		1282	ATCAACTCTCCCAACAGAGACTACACATCTGTGAGCAGTACAGCGCACCGAGGCGCGC	1341		
Oy		221	HtSHtSLeuPheLeu	225		
Db		1342	CACCACTGTCTCTG	1356		
<b>RESULT 32</b>						
ID	ACC44640					
ACC44640	standard; DNA; 4692 BP.					
XX	AC	ACC44640;				
XX	DT	29-MAY-2003 (first entry)				
DE	Vector pDareRedINI nucleotide sequence SEQ ID NO:29.					
KX	Chromosome-based platform; artificial chromosome; eukaryotic chromosome;					
KW	act site; integrase; recombinase; Acs; gene therapy; transgenic animal;					
KM	platform artificial chromosome expression system; gene; ds.					
XX	OS	DiscoSoma sp.				
OS	Synthetic.					
PX	PN	WO200297059-A2.				
XX	PD	05-DEC-2002.				
PF	30-MAY-2002;	2002WO-US017452.				
PR	30-MAY-2001;	2001US-0294758P.				
XX	21-MAR-2002;	2002US-0366891P.				
PA	(CHRO-) CHROMOS MOLECULAR SYSTEMS INC.					
XX	Perkine E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E,					
PI	Stewart S, Shellard J,					

XX		WP1; 2003-140461/13.
DR		
PT	Novel eukaryotic chromosome comprising one or many att sites which permit site-directed integration in the presence of lambda-integrase, useful for site-specific recombination-directed integration of DNA of interest.	
PR		
XX		
PS	Example 1; Page 182-184; 272pp; English.	
XX		
CC	The present invention describes a eukaryotic chromosome (I) comprising one or several att sites, where an att site is heterologous to the chromosome, and permits site-directed integration in the presence of lambda-integrase. Also described: (1) a platform artificial chromosome expression system (Acce) (II) comprising several sites that participate in recombinase catalyzed recombination; and (2) a method (M1). For introducing a heterologous nucleic acid into a platform artificial chromosome. (I) can be used in gene therapy. (M1) is useful for introducing a heterologous nucleic acid molecule into a platform artificial chromosome, preferably an Acce. (II) is useful for producing a transgenic animal (e.g., a fish, insect, reptile, amphibian, arachnid, or mammal) by introducing (II) by cell fusion, lipid-mediated transfection by a carrier system, microinjection, microcell fusion, electroporation, microparticle bombardment or direct DNA transfer into an embryonic cell, preferably a stem cell or an embryo. (II) comprises a heterologous nucleic acid that encodes a therapeutic product which is useful for making a library of Acces comprising random portions of a genome. ACC44612 to ACC44732 and ABP96550 to ABP96557 represent sequences used in the exemplification of the present invention	
CC		
CC		
CC		
XX	Sequence 4692 BP, 1111 A; 1268 C; 1247 G; 1066 T; 0 U; 0 Other;	
SQ		
	Alignment Scores:	
	Pred. No.: 1,1e-138 Length: 4692	
	Score: 1210.00 Matches: 224	
	Percent Similarity: 100.00% Mismatch: 1	
	Best Local Similarity: 99.56% Mismatches: 0	
	Query Match: 99.67% Indels: 0	
	Gaps: 0	
DB:		
US-10-006-922A-12 (1-225) X ACC44640 (1-4692)		
OY	1 MecATGSSerSerLyABsnValILlelygBilPhemetrApheLyvAlArqMetGlucly 20 :::	
Db	682 GTGCCTCCTCCAAGAAACGTCAATCAAGAAGATTACGCCCTTCAAAGTGCGCATGAAGGC 741	
OY	21 ThrValAMngLYHIAsgLpNheglunlleglnclYglunclYglunclYARqProTYrGLunclY 40	
Db	742 ACCGGAAACGGCACGAAGTTGAGATTCAGAGGCGAGGGCGCGCCCTTAGAGAGGC 801	
OY	41 HISAbnTrrrValLYLeulYvAlThrrYsgLYglYProLeuPProPheAlATrAspIlle 60 :::	
Db	802 CAACAACCCGTGAAGCTTAAGGTACCAAGGGCGGCCCTCGCCTTGCGGAGACTC 861	
OY	61 LeuSerProginPnheglInTyrgLYserLYsValYrVAllySHIpcoAlAspIllePro 80 :::	
Db	869 CTGTCCCCCAAGTTCCAGTACGGCTCCCAAGGTGTACGTGAAGCAACC GCCGACATCCCC 921	
OY	81 ApplYrlyblYbLeuSerPheProglunclYphelysrTrpglWAgYValMeAsnPheglu 100 :::	
Db	922 GACTCAAGAAAGCTGTCTCTCCGCCAAGGGCTTCAAGTGGAGCCGGTGAATAACTTCAG 981	
OY	101 AAPGAYglYvalValThThrValThrgLnAspSerserLeugLnAspglYCpPheilleTy 120 :::	
Db	982 GACGGCGGGGTGTACCGGTACCCAGAACCTCTCCCTCGACGAGACGGCTGTTATCTAC 1041	
OY	121 LySVallySpheilleglYValASnpheProSerAppglYProValMetGlnLYslYThr 140 :::	
Db	1042 AAGGTGAAGTTTCATCGCGGTGAATCTTCCCCTCCGACGGCCCCGTATATCAAGAAAGAC 1101	
OY	141 MetGlyTrpglUAlAsErThrGlUAgrleuTYrProAgaarpglYAlleulYglYglu 160 :::	
Db	1102 ATGGGCTGGAGGCTCTCCACCGAGGGCTGTAACTCCCGCAGCGGTGTAAAGGGCGAG 1161	





AAD09979/C  
ID AAD09979 standard; DNA; 7910 BP.  
XX AC AAD09979;  
XX DT 12-SEP-2001 (first entry)  
DE pB1T(dHSP)-RFP-oXoXS/BH construct for transfection of pacific oysters.  
XX KM Simian virus 40, SV40; fertility; reproduction; gametogenesis; fruitfly;  
XX RM microinjection; infection; plasmid pB1T(dHSP)-RFP-oXoXS/BH; oyster; HSP;  
XX KW embryogenesis; RFP; heat shock promoter; red fluorescent protein; ds.  
XX OS Crassostrea gigas.  
OS Rheus macaque polyoma virus.  
OS Drosophila melanogaster.  
OS Unidentified.  
OS Chimeric.  
XX MO200148224-A1.  
PN XX 05-JUL-2001.  
PD XX 22-DEC-2000; 2000WO-AU001596.  
PF XX 24-DEC-1999; 99AU-00004884.  
PR XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA XX Threshner R, Hinds L, Hardy C, Whyard S, Vignarajan S, Grewe PW;  
PI Pl Pacil J;  
XX DR WPI; 2001-425672/45.  
DR PT Novel construct for preventing embryogenesis in animals comprises native  
PT promoter, blocking DNA which abrogates function of crucial gene and  
PT genetic switch to regulate expression/repression of blocker/gene  
PT knockout.  
PS Example 8; Page 193-196; 241pp; English.  
XX CC The invention relates to a construct which allows animals to be bred in  
CC captivity but renders them infertile in the wild by allowing reversible  
CC control over fertility and reproduction. The construct comprises a native  
CC promoter, a blocking DNA sequence contoured for and designed to abrogate  
CC a crucial gene's function or to cause its mis-expression, and a genetic  
CC switch to regulate controlled expression/repression of the blocker/gene  
CC knockout. The construct is useful for preventing embryogenesis or  
CC gametogenesis in animals by stably transforming an animal cell with the  
CC construct by microinjection, transfection or infection, where the  
CC construct stably integrates into the genome by homologous recombination,  
CC and implanting the cell into a host organism, where a whole animal  
CC develops from the implanted cell. The present DNA sequence is plasmid  
CC construct pB1T(dHSP)-RFP-oXoXS/BH used for transfection of Pacific  
CC oysters. The plasmid comprises oyster Hox gene under the under the  
CC control of Drosophila heat shock promoter (dHSP), red fluorescent protein  
CC and SV40 Polya

XX SQ Sequence 7910 BP; 2016 A; 1893 C; 1924 G; 2077 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,3e-138 Length: 7910  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 4 Gaps: 0

US-10-006-9922A-12 (1-225) x AAD09979 (1-7910)

OY 1 MetaAggSerIyAsnValIlelysgLUPheMetArghelysValArgMctGlunclly 20  
:::  
DB 2705 GGCGCCTCTCCAMAGACGTCAAGAGTTTCATGCCGTTTAAGAGTGCGCATGGAGGC 2646

Oy		21	ThrValIamGlyYhiSGIupheGluIlleglUGlyVGIugLyglUGlyAargProTyrgUgLy	40
Db		2645	ACCGTGAACGGCCACGAGTTCTGAGATCGAGGGGCGAGGGCCGCCCTTCAGAAAGGC	2586
Oy		41	HisaenThrValIysLeuLysValIThrLysGlyVGlyProLeuProPhealaATrpAspIle	60
Db		2585	CACAACACCGTGAAAGCTGAAGGTGACCAAGGCGGGCCCCCTGCCCTTCGCTGGGAATC	2526
Oy		61	LeuSerProGlnpheGlnTyrgYserLysValITyrValIysHisProalaaSpIlePro	80
Db		2525	CTGTCGCCCGCATGTTCCAGTACGGGCTCAAGGTGACTGTGAAGCACCCCGCGCATCCCC	2466
Oy		81	AspTyrILysLysLeuSerPheProGluGLYPhelytTrpGluArgValIMetAsnPheGlu	100
Db		2465	GACTACAGAAGCGTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTAATGAATTGGAG	2406
Oy		101	AspGlyGlyValValIThrValIThrGlnAspSerserLeuGlnaspGlyCySPheileTy	120
Db		2405	GACGGCGCGGTGTGTGACCGTGAACCGAGACTCCTCGACAGACGGCTGCTTCATCTAC	2346
Oy		121	LysValIysPheIIeGlyValaAsnphEProSerAspGlyProvalMetGlnLysLysThr	140
Db		2345	AAGGTGAAGTTCACTGGCGGTGAACTTCCCTCCGACGGCCCCGTATGCAAGAAAGACC	2286
Oy		141	MetylITrpGluIasertHngIuarGLyTyrProArgAspGlyValIleLysGlyGU	160
Db		2285	ATGGAGCTGGAGGGGCTTCACCGAGCGCTGTACCCCGCCGACGGCGTGTGAGGGCGAG	2226
Oy		161	IlehialysalaleuLYsLeuLYsAspGlyGlyHIstYrLeuValIGluPheLysSerIle	180
Db		2225	ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACTCGGTGGAGTTCAATGCATCC	2166
Oy		181	TyrMetalalybLYsProvalGlnLeuProGlyTYrtYrtyrValIaaseryLeuasp	200
Db		2165	TACATGGCCAAGAACCCCGTGCACCTGCCGGCTACTACTACGTGACTCAACCTGGAC	2106
Oy		201	IllethrSerHisangIuaAspyIThrIleValIGluGlnTyrgIuarGTThrGluGLYArg	220
Db		2105	ATCACCTCCCACAACGAGACTACACACTGTGAGAGACTACGAGCCGACCGAGGGCGGC	2046
Oy		221	HisHisLeuPheLeu	225
Db		2045	CACCACTGTGCTCTG	2031
RESULT 36				
ABS56664	ID	ABS56664	standard; DNA; 9320 BP.	
XX	AC	ABS56664;		
XX	XX			
DT	24-JAN-2003	(first entry)		
DE	Plasmid p5-Puro-CMV-(N'-EGFP)-CMV-Red (EGFP-EJ) DNA.			
XX				
KW	Vector; genotoxicity; food component; cosmetic; medicine; cancer;			
KW	genetic stability; hygromycin; ds.			
XX				
OS	Synthetic.			
XX				
FH	Key	Location/Qualifiers		
FT	misc_feature	1..1592		
FT		/tag= a		
FT		/note= "retroviral vector p5NM"		
FT	CDS	1617..2216		
FT		/tag= b		
FT		/product= "puromycin resistance gene from pRetron"		
FT	Promoter	2267..2848		
FT		/tag= c		
FT		/note= "CMV promoter from pEGFP-N1"		
PT	CDS	2906..3348		
FT		/tag= d		
FT		/product= "N'-EGFP"		



FT misc\_feature 3374. .3392  
 FT /tag= e  
 FT /note= "retroviral vector p5NM"  
 FT promoter 3411. .3392  
 FT /tag= f  
 FT /note= "CMV promoter from pDered1-N1"  
 FT CDS 4038. .4718  
 FT /tag= g  
 FT /product= "pDered1-N1"  
 FT CDS 4766. .5508  
 FT /tag= h  
 FT /product= "EGFP-EJ"  
 FT misc\_feature 5527. .9320  
 FT /tag= 1  
 FT /note= "retroviral vector p5NM"  
 FT  
 FT WO20020740-A2.  
 PD 12-SEP-2002.  
 PD 28-FEB-2002; 2002WO-EP002194.  
 PF 05-MAR-2001; 2001DE-01010449.  
 PR (WIES/) WISMUELLER L.  
 PA  
 XX  
 PI Wismueller L.  
 XX WPI; 2002-682909/73.  
 XX  
 PT New vector containing recombination sequences, useful e.g. for assessing  
 PT compounds for genotoxicity, cancer susceptibility and mutational  
 PT frequency.  
 PT  
 PS Disclosure; Page 104-106; 106pp; German.  
 XX  
 CC This invention describes a novel vector containing, at least, two  
 CC homologous DNA sequences that, through DNA exchange, induce a change in a  
 CC gene present in the vector. The products of the invention are used to  
 CC assess genotoxicity of compounds (e.g. components of foods, cosmetics and  
 CC medicines), to determine susceptibility for development or progression of  
 CC cancer, for analysis of (non-)conservative recombination processes, and  
 CC joining and inactivating/reverting mutations, for characterization of  
 CC selected genes with respect to maintenance or reduction of genetic  
 CC stability, and to determine genetic (in)stability and recombinational or  
 CC mutational frequency of cells, tissues or organisms. The vector allows an  
 CC analysis to be performed in hours (making it suitable for large-scale,  
 CC routine use), and can be used with a variety of eukaryotic cells and  
 CC animals. All known types of DNA recombinations can be examined. This  
 CC sequence represents a plasmid p5-Puro-CMV-(N'-EGFP)-CMV-Red(EGFP-EJ) DNA,  
 CC described in the disclosure of the invention  
 CC  
 SQ Sequence 9320 BP; 2124 A; 2687 C; 2474 G; 2035 T; 0 U; 0 Other;  
 XX  
 XX  
 Alignment Scores:  
 Pred. No.: 2.89e-138 Length: 9320  
 Score: 1210.00 Matches: 224  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.56% Mismatches: 0  
 Query Match: 99.67% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-006-922a-12 (1-225) x ABS5664 (1-9320)  
 QY 1 MetAAGSerSerLyAsnValIleLySGUbuPheMetArgPheLyValArgMetGluGly 20  
 DB 4041 GGGCCCTCTCCAGAAAGCTATCAAGAGTTCAAGCTTCAAGGTCGCAATGAGGAGC 4100  
 QY 21 ThrValAenGlyVHISglupheGluIleGluGlyGluGlyVArgpProTyrgluGly 40  
 DB 4101 ACCGCGAACGCGCACAGTTCCAGATCGAGGCGAGGCGGAGGCGCCCTTACAGAGGC 4160  
 QY 41 HIsaenThrValLyLeuLyValThrLySGlyGlyProLeuProPheAlaTrpaspIle 60

DB 4161 CACACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCGCTTCGCTGGAGATC 4220  
 QY 61 LeuSerProGlnPheGlnTyrglySerLyValTyrrVallyshisProlAaspIlePro 80  
 DB 4221 CTGTCCCCCAGTTCCAGTACGGGCTCCAGGTGTAAGTGAAGCACCCCGCGCATCCTCC 4280  
 QY 81 AspTyrLyshLyLeuSerPheProGluGlyPheLyThrGluArgValMetAsnPheGlu 100  
 DB 4281 GACTTACAAAGAGCTGCTCTTCCCGAGGGCTTCAAGTGGAGACCGCGTGAATGAACTTGAG 4340  
 QY 101 AspGlyLyValValThrValThrGlnAaspSerSerLeuGlnAaspGlyCyPheIleTyr 120  
 DB 4341 GACGCGCGCGTGTGACCGTGAACCCAGAGACTCTCCCTGACAGACGGCTGCTTCACTTAC 4400  
 QY 121 LyValLyshPheIleGlyValAaspPheProSerAaspGlyProValMetGlnLyshThr 140  
 DB 4401 AAGGTGAAGTTCACTCGGCGTGAAGTCTCCCTCCGACGGCCCGTAAATGCAAGAAAGACC 4460  
 QY 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAaspGlyValLeuLyshGly 160  
 DB 4461 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGCGCTGTGAAGGCGAG 4520  
 QY 161 ILeHisLyshAlaLeuLyshLeuLyAaspGlyGlyHisTyrLeuValGluPheLyshIle 180  
 DB 4521 ATCCAAAGGCGCTGAAAGCTGAAGACGGCGGCCACTACCTGTGTGAAGTTCAAGTTCATC 4580  
 QY 181 TyrMetAlaLyshLyProValGlnLeuProGlyTyrTyrTyrValAaspSerLyshLeuA 200  
 DB 4581 TACATGGCAAGAGCCCGTGACCTGCGCTACTACTACCTGAGTCAAGTCAAGTGGAC 4640  
 QY 201 IleThrSerHisAenGluAaspTyrThrIleValGluGlnTyrGluArgThrGluGly 220  
 DB 4641 ATCACTCTCCCAACGAGACTACCACTATCGTGAAGAGTACGAGCGCACCGAGGCGCG 4700  
 QY 221 HIsHisLeuPheLeu 225  
 DB 4701 CACCACTGTCTCTG 4715  
 RESULT 37  
 AD104104  
 ID AD104104 standard; DNA; 9658 BP.  
 XX  
 AC AD104104;  
 XX  
 DT 22-APR-2004 (first entry)  
 XX  
 DE Vector pTnMod (Oval/Red) - Qual.  
 XX  
 XX ss; vector; transposase; promoter; insertion sequence; kozak; Tn10;  
 KW ovalbumin; vitellogenin; polyA; conalbumin; enhancer; ovomucoid;  
 KW signal sequence; transgenic animal; Huntington's disease;  
 KW alpha-1-antitrypsin deficiency; Alzheimer's disease; breast cancer;  
 KW cystic fibrosis; galactosaemia; congenital hypothyroidism;  
 KW maple syrup urine disease; neurofibromatosis; phenylketonuria;  
 KW sickle cell disease; Smith-Lemli-Opitz Syndrome; autoimmune disease;  
 KW shipping fever; cattle; maelitis; bacterial; viral; infection; Types I;  
 type; II; diabetes.  
 OS Synthetic.  
 FH  
 FH Key Location/Qualifiers  
 FT 1. .130  
 FT misc\_feature  
 FT /tag= a  
 FT /note= "Remainder of Fl(-) on from pBluescriptII sk(-)"  
 FT  
 FT misc\_feature  
 FT /tag= b  
 FT /note= "Residue from ligation of restriction enzyme sites  
 FT used in construction"  
 FT  
 FT promoter  
 FT 133. .1777  
 FT /tag= c  
 FT /note= "CMV promoter/enhancer from pGwiz"  
 FT  
 FT misc\_feature  
 FT 1778. .1779

```

FT      /*tag= d
FT      /note= "Residue from ligation of restriction enzyme sites
FT      used in construction"
FT      CDS
FT      1780..2987
FT      /*tag= e
FT      /product= "Transposase"
FT      /note= "Modified from Tn10 by optimising codons for
FT      stability"
FT      2988..2993
FT      /*tag= f
FT      /note= "Two engineered stop codons"
FT      2995..3410
FT      /*tag= g
FT      /note= "Taken from pGWiz"
FT      3415..3718
FT      /*tag= h
FT      /note= "Non-coding DNA residual from pNK2859"
FT      3719..3761
FT      /*tag= i
FT      /note= "Non-coding lambda DNA residual from pNK2859"
FT      3762..3831
FT      /*tag= j
FT      /note= "Left insertion sequence recognised by Tn10"
FT      3832..3837
FT      /*tag= k
FT      /note= "Residue from ligation of restriction enzyme sites
FT      used in construction"
FT      3838..4044
FT      /*tag= l
FT      /note= "Multiple cloning site from pBluescript1 sk(-)"
FT      4045..4049
FT      /*tag= m
FT      /note= "Residue from ligation of restriction enzyme sites
FT      used in construction"
FT      4050..4934
FT      /*tag= n
FT      /note= "Japanese quail ovalbumin promoter"
FT      /note= "Includes steroid-dependant response element"
FT      4935..4942
FT      /*tag= o
FT      /note= "Residue from ligation of restriction enzyme sites
FT      used in construction"
FT      4943..5092
FT      /*tag= p
FT      /note= "Japanese quail ovalbumin signal sequence"
FT      5093..5098
FT      /*tag= q
FT      /note= "Residue from ligation of restriction enzyme sites
FT      used in construction"
FT      5099..5591
FT      /*tag= r
FT      /note= "Dated reporter coding sequence including polyA
FT      from pDREd1.1"
FT      5592..5597
FT      /*tag= s
FT      /note= "Residue from ligation of restriction enzyme sites
FT      used in construction"
FT      5598..6036
FT      /*tag= t
FT      /note= "Multiple cloning site from ZeroBlunt Topo cloning
FT      vector"
FT      6037..6042
FT      /*tag= u
FT      /note= "Residue from ligation of restriction enzyme sites
FT      used in construction"
FT      6043..6475
FT      /*tag= v
FT      /note= "Multiple cloning site from pBluescript1 sk(-)"
FT      6476..6480
FT      /*tag= w
FT      /note= "Residue from ligation of restriction enzyme sites
FT      used in construction"
FT      6481..6550
FT      insertion_seq

```

```

FT      /*tag= x
FT      /note= "Right insertion sequence recognised by Tn10"
FT      misc_feature
FT      6551..6592
FT      /*tag= y
FT      /note= "Non-coding lambda DNA residual from pNK2859"
FT      misc_feature
FT      6593..7457
FT      /*tag= z
FT      /note= "Non-coding DNA residual from pNK2859"
FT      misc_feature
FT      7458..9658
FT      /*tag= aa
FT      /note= "Remainder from pBluescript1 sk(-)"
FT      PN
FT      WO2004003157-A2.
FT      XX
FT      PD
FT      08-JAN-2004.
FT      XX
FT      PF
FT      26-JUN-2003; 2003WO-US020389.
FT      XX
FT      PR
FT      26-JUN-2002; 2002US-0392415P.
FT      PR
FT      21-JAN-2003; 2003US-0441377P.
FT      PR
FT      21-JAN-2003; 2003US-0441381P.
FT      PR
FT      21-JAN-2003; 2003US-0441392P.
FT      PR
FT      21-JAN-2003; 2003US-0441405P.
FT      PR
FT      21-JAN-2003; 2003US-0441447P.
FT      PR
FT      21-JAN-2003; 2003US-0441502P.
FT      XX
FT      PA
FT      (TRAN-) TRANSGENRES LLC.
FT      PA
FT      (LOU) UNIV LOUISIANA STATE & AGRIC & MECH COLL.
FT      PI
FT      Cooper RK, Cadd GG, Floretti WC, Deboer KF;
FT      WPI; 2004-083042/08.
FT      XX
FT      PR
FT      New vector comprising a transposase gene operably linked to a promoter,
FT      useful in treating Alzheimer's disease, breast cancer, phenylketonuria,
FT      autoimmune diseases, bacterial or viral infections or Types I and II
FT      diabetes.
FT      XX
FT      PS
FT      Example 4; SEQ ID NO 4; 150bp; English.
FT      XX
FT      CC
FT      This sequence represents a new vector, pTnMod, comprising a transposase
FT      gene operably linked to a first promoter and one or more gene of interest
FT      operably linked to one or more additional promoters, where the one or
FT      more genes of interest and their operably linked promoters are flanked by
FT      CC
FT      transposase insertion sequences recognized by the transposase and the
FT      CC
FT      first promoter comprises a modified Kozak sequence comprising ACCATG. The
FT      CC
FT      transposase is a Tn10 transposase and is modified in one to twenty of the
FT      CC
FT      first codons. The gene of interest in operably linked to a second
FT      CC
FT      promoter, which is a constitutive promoter or an inducible promoter, e.g.
FT      CC
FT      an ovalbumin or a vitellogenin promoter. The vector further comprises a
FT      CC
FT      polyA sequence, a conalbumin polyA sequence, operably linked to the
FT      CC
FT      transposase gene. The vector further comprises two stop codons operably
FT      CC
FT      linked to the transposase gene. The first or second gene of interest is
FT      CC
FT      operably linked to a second or third promoter, respectively. The first
FT      CC
FT      and a second gene of interest are operably linked to a second promoter.
FT      CC
FT      The vector further comprises an enhancer operably linked to the one or
FT      CC
FT      more genes of interest. The enhancer comprises at least a portion of an
FT      CC
FT      ovalbumin enhancer. The vector further comprises an egg chitinogen
FT      CC
FT      targeting sequence, operably linked to the one or more genes of interest.
FT      CC
FT      The vector is useful in producing transgenic animals with desired
FT      CC
FT      proteins or molecules. They are also useful in treating Huntington's
FT      CC
FT      disease, alpha-1-antitrypsin deficiency, Alzheimer's disease, breast
FT      CC
FT      cancer, cystic fibrosis, galactosaemia, congenital hypothyroidism, maple
FT      CC
FT      syrup urine disease, neurofibromatosis, phenylketonuria, sickle cell
FT      CC
FT      disease, Smith-Lemli-Opitz Syndrome, autoimmune diseases, shipping fever
FT      CC
FT      in cattle, mastitis, bacterial or viral infections or Types I and II
FT      CC
FT      diabetes.
FT      CC
FT      SQ
FT      Sequence 9658 BP; 2572 A; 2397 C; 2148 G; 2541 T; 0 U; 0 Other;
FT      Alignment Scores:
FT      Pred. No.: 3.04e-138 Length: 9658

```

Score: 1210.00 Matches: 224  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.56% Mismatches: 0  
 Query Match: 99.67% Indels: 0  
 DB: 12 Gaps: 0

US-10-006-922a-12 (1-225) x AD104104 (1-9658)

QY 1 MetATGSeSerSeLyASnValIleLyGluPhneMetArGPhelyeValAArgMetGluGly 20  
 DB 5099 GTGGCCCTCTCCAAAGACTCTATCAAGAGTTCAAGCTTCAAGTGGCATGAGAGGC 5158  
 QY 21 ThrValaangLYHIGluPhneGluIleGluGlyGluGlyArgProTyrGluGly 40  
 DB 5159 ACCGTGAAGCGCCAGAGTTGAGATCGAGGCGGAGGCGGAGGCGCCCTTACGAGGAGC 5218  
 QY 41 HIsaAntThValLyLeuLyLeuValThrLyGlyGlyProLeuProPhaLaATrPaaple 60  
 DB 5219 CACAAACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTTCCCTGAGGAGATC 5278  
 QY 61 LeuSeArProGlnPhneGlnTyrGlySerLySValTyrValLyHIsProLaAspIlePro 80  
 DB 5279 CTGTCCCCCACTTCCAGTACGGCTCCAGGTGATGAAACACCCCGCAGCATCCCC 5338  
 QY 81 AepTyrLyLyLeuSeArProGluGlyPheLySTrPGLuArgValMeCaAnPheGlu 100  
 DB 5339 GACTACAAAGAGCTGTCTTCCCGAGGCGCTTCAAGTGGAGACGCGTGAATGACTTCGAG 5398  
 QY 101 AepGlyGlyValValThrValThrGlnAAserSerLeuGlnAAspGlyVcyPheHeIleTyr 120  
 DB 5399 GACGGCGGGGTGTGACCGGTGACCCAGACTCTCTCTCAAGACGCGCTGTCACTGAC 5458  
 QY 121 LySValLySPhelIleGlyValAAsPheProSeArSpGlyProValMeGlnLySlyThr 140  
 DB 5459 AAGGTGAAGTTCAATCGCGCTGAATTCCTCCCTCCAGCGGCCCGTATGCAAGAAAGACC 5518  
 QY 141 MetGlyTyrGluAAserThrgluArgLeuTyrProArgAAspGlyValLeuLyGlyIle 160  
 DB 5519 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGAGCGCGTGTGAGAAAGGCGAG 5578  
 QY 161 HIsaLyLyAAlaLeuLyLeuLySAspGlyGlyHIsTyrLeuValGluPheLySserIle 180  
 DB 5579 ATCCCAAGAGCCCTGAAGCTGAAGAGCGCGGCACTACTGTGAGTCAAGTCCATC 5638  
 QY 181 TyrMetAlaLySlySProValGlnLeuProGlyTyrTyrTyrValAAsPserLySLeuAAsp 200  
 DB 5639 TACATGGCCAAAGAGCCCTGACGTGCGCGGTACTTACTAGTGAATCCAAAGCTGGAGC 5698  
 QY 201 ILeThSerHIsaAngLyAAsPlyTyrHrIleValGluGlnTyrGluArgThrgluGlyArg 220  
 DB 5699 ATCACTCCCAAGAGGACTACACATCGTGAAGACAGTACAGAGCGCACCGAGGCGCGC 5758  
 QY 221 HIsaLyLeuPheLeu 225  
 DB 5759 CACCACTGTCTCTG 5773

RESULT 38

AD104103 standard; DNA; 9678 BP.

AD104103;

22-APR-2004 (first entry)

Vector pTnMod (Oval/Red)-Chicken.

ss; vector; transposase; promoter; insertion sequence; Kozak; Tn10;  
 ovalbumin; vitellogenin; polyA; conalbumin; enhancer; ovomucoid;  
 signal sequence; transgenic animal; Huntington's disease;  
 alpha-1-antitrypsin deficiency; Alzheimer's disease; breast cancer;  
 cystic fibrosis; galactosemia; congenital hypothyroidism;  
 maple syrup urine disease; neurofibromatosis; phenylketonuria;  
 sickle cell disease; Smith-Lemli-Opitz Syndrome; autoimmune disease;

KW shipping fever; cattle; mastitis; bacterial; viral; infection; Types I;  
 KW type; II; diabetes.  
 XX Synthetic.  
 OS  
 XX  
 XX  
 FH Key  
 FT misc\_feature  
 FT 1..130  
 /tag= a  
 /note= "Remainder of Pl(-) on from pBluescriptII SK(-)"  
 FT 131..132  
 /tag= b  
 /note= "Residue from ligation of restriction enzyme sites  
 used in construction"  
 FT 133..1777  
 /tag= c  
 /note= "CMV promoter/enhancer from pGWiZ"  
 FT 1778..1779  
 /tag= d  
 /note= "Residue from ligation of restriction enzyme sites  
 used in construction"  
 FT 1780..2987  
 /tag= e  
 /note= "Transposase"  
 FT 2988..2993  
 /tag= f  
 /note= "Modified from Tn10 by optimising codons for  
 stability"  
 FT 2994  
 /tag= g  
 /note= "Two engineered stop codons"  
 FT 2995..3410  
 /tag= h  
 /note= "Taken from pGWiZ"  
 FT 3411..3718  
 /tag= i  
 /note= "Non-coding DNA residual from pNK2859"  
 FT 3719..3761  
 /tag= j  
 /note= "Non-coding lambda DNA residual from pNK2859"  
 FT 3762..3831  
 /tag= k  
 /note= "Left insertion sequence recognised by Tn10"  
 FT 3832..3837  
 /tag= l  
 /note= "Residue from ligation of restriction enzyme sites  
 used in construction"  
 FT 3838..4044  
 /tag= m  
 /note= "Multiple cloning site from pBluescriptII SK(-)"  
 FT 4045  
 /note= "May be used to insert the sequence of interest  
 into the vector"  
 FT 4046  
 /tag= n  
 /note= "Residue from ligation of restriction enzyme sites  
 used in construction"  
 FT 4050..4951  
 /tag= o  
 /note= "Upstream elements"  
 FT 4952..5112  
 /tag= p  
 /note= "Chicken ovalbumin signal sequence"  
 FT 5113..5118  
 /tag= q  
 /note= "Residue from ligation of restriction enzyme sites  
 used in construction"  
 FT 5119..6011  
 /tag= r  
 /note= "Red reporter coding sequence including polyA  
 from pDared1.1"

```

FT misc_feature 6012..6017
FT /*tag= 8
FT /note= "Residue from ligation of restriction enzyme sites
FT used in construction"
FT misc_signal 6018..6056
FT /*tag= t
FT /note= "Multiple cloning site from ZeroBlunt Topo cloning
FT vector"
FT misc_signal 6057..6062
FT /*tag= u
FT /note= "Residue from ligation of restriction enzyme sites
FT used in construction"
FT misc_signal 6063..6435
FT /*tag= v
FT /note= "Multiple cloning site from pBluescriptII sk(-)"
FT misc_signal 6436..6500
FT /*tag= w
FT /note= "Residue from ligation of restriction enzyme sites
FT used in construction"
FT insertion_seq 6501..6570
FT /*tag= x
FT /note= "Right insertion sequence recognised by Tn10"
FT misc_feature 6571..6612
FT /*tag= y
FT /note= "Non-coding lambda DNA residual from pNK2859"
FT misc_feature 6613..7477
FT /*tag= z
FT /note= "Non-coding DNA residual from pNK2859"
FT misc_feature 7478..9678
FT /*tag= aa
FT /note= "Remainder from pBluescriptII sk(-)"
FT
FT WO2004003157-A2.
FT
FT 08-JAN-2004.
FT
FT 26-JUN-2003; 2003WO-US020389.
FT
FT 26-JUN-2002; 2002US-0392415P.
FT 21-JAN-2003; 2003US-0441377P.
FT 21-JAN-2003; 2003US-0441381P.
FT 21-JAN-2003; 2003US-0441392P.
FT 21-JAN-2003; 2003US-0441405P.
FT 21-JAN-2003; 2003US-0441447P.
FT 21-JAN-2003; 2003US-0441502P.
FT
FT (TRAN-) TRANSGENRES LLC.
FT (LOU) UNIV LOUISIANA STATE & AGRIC & MECH COLL.
FT
FT Cooper RK, Cadd GG, Fioretto WC, Deboer KF;
FT
FT WPI; 2004-083042/08.
FT
FT New vector comprising a transposase gene operably linked to a promoter,
FT useful in treating Alzheimer's disease, breast cancer, phenylketonuria,
FT autoimmune diseases, bacterial or viral infections or Types I and II
FT diabetes.
FT
FT Example 3; SEQ ID NO 3; 150bp; English.
FT
XX This sequence represents a new vector, pTnMod (Oval/Red)-Chicken,
XX comprising a transposase gene operably linked to a first promoter and one
XX or more gene of interest operably linked to one or more additional
XX promoters, where the one or more genes of interest and their operably
XX linked promoters are flanked by transposase insertion sequences
XX recognized by the transposase and the first promoter comprises a modified
XX korak sequence comprising ACCATG. The transposase is a Tn10 transposase
XX and is modified in one to twenty of the first codons. The gene of
XX interest is operably linked to a second promoter, which is a constitutive
XX promoter or an inducible promoter, e.g. an ovalbumin or a vitellogenin
XX promoter. The vector further comprises a polyA sequence, a consalumin
XX polyA sequence, operably linked to the transposase gene. The vector
XX further comprises two stop codons operably linked to the transposase

```

```

CC gene. The first or second gene of interest is operably linked to a second
CC or third promoter, respectively. The first and a second gene of interest
CC are operably linked to a second promoter. The vector further comprises an
CC enhancer operably linked to the one or more genes of interest. The
CC enhancer comprises at least a portion of an ovalbumin enhancer. The
CC vector further comprises an egg directing sequence, ovalbumin or
CC ovomucoid signal sequence or vitellogenin targeting sequence, operably
CC linked to the one or more genes of interest. The vector is useful in
CC producing transgenic animals with desired proteins or molecules. They are
CC also useful in treating Huntington's disease, breast cancer, cystic fibrosis,
CC deficiency, Alzheimer's disease, phenylketonuria, sickle cell disease, Smith-Lemli-
CC galactosaemia, congenital hypothyroidism, maple syrup urine disease,
CC neurofibromatosis, phenylketonuria, sickle cell disease, Smith-Lemli-
CC Optiz Syndrome, autoimmune diseases, shipping fever in cattle, mastitis,
CC bacterial or viral infections or Types I and II diabetes.
CC
XX
XX
SQ Sequence 9678 bp; 2586 A; 2396 C; 2152 G; 2544 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3,05e-138 Length: 9678
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: Gaps: 0
US-10-006-922a-12 (1-225) x AD104103 (1-9678)
QY 1 MetArgSerSerIysAsnValIleYlgVluPheMetArgPheIysValArgMetGluGly 20
DB ::::
5119 GTGGGCTCTCCAAAGACGTCATCAAGAGTATGCGCTTCAAGGCGCCATGAGAGGC 5178
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyValGluGlyAspProIyrGluGly 40
DB |||||
5179 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGGCGAGGGCCGCCCTTACGAGAGGC 5238
QY 41 HisAsnThrValIysLeuIysValThrValGlyIyrProLeuProPheAlaTrpAspIle 60
DB |||||
5239 CACACACCGTGAAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGGCTGGAGATC 5298
QY 61 LeuSerProGlnPheGlnIyrGlySerIysValIyrValIysHisProAlaAspIlePro 80
DB |||||
5299 CTGTCCCCCAGTTCCTCGTACGGCTCCCAAGGTGATCGAAGCACCCCGCCGACATCCC 5358
QY 81 AspTyrIyrIysLeuSerPheProGluGlyPheIyrTrpGluArgValMetAsnPheGlu 100
DB |||||
5359 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCGAG 5418
QY 101 AspGlyIyrValIyrThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB |||||
5419 GACGGCGGCTGTGTGACCGTGACCCAGGACTCTCTCCCTGCAGAGCGGCTGTTCTAC 5478
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIleThr 140
DB |||||
5479 AAGGTGAAGTTCACTGGCGGTGAACCTTCCCTCGACGCCCGCTGATGACGAGAAAGACC 5538
QY 141 MetGlyIyrTrpGluIysSerThrGluArgLeuTyrProArgAspGlyValIleLysGly 160
DB |||||
5539 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGGTGAAAGGGCGAG 5598
QY 161 IleHisIysAlaLeuIysLeuIysAspGlyIyrHisIyrLeuValGluPheIysSerIle 180
DB |||||
5599 ATCCACAAAGGCCCTGAAGCTGAAGAGAGGGGCCACATCACTGTGTGAATCAATCCATC 5658
QY 181 TyrMetAlaIysIysProValGlnLeuProGlyIyrTyrIyrValAspSerIysLeuAsp 200
DB |||||
5659 TACATGGCCAAAGAGCCCGTGCAGCTGCCGGCTACTACTACGTGAGACTCCAAAGCTGAC 5718
QY 201 IleThrSerHisAsnGluAspIyrThrIleValGluIleIyrGluArgThrGluGlyArg 220
DB |||||
5719 ATCACTCTCCCAACAGCGAGACTACACATCGTGGAGATACGAGCGACCGAGAGGGCGGC 5778
QY 221 HisHisLeuPheLeu 225

```

Db 5779 CACCACCTGTTCTG 5793

RESULT 39

AD104102 standard; DNA; 10263 BP.

AD104102;

22-APR-2004 (first entry)

Vector pTmMod (CMV/Red).

as; vector; transposase; promoter; insertion sequence; Kozak; Tn10; ovalbumin; vitellogenin; polyA; conalbumin; enhancer; ovomucoid; signal sequence; transgenic animal; Huntington's disease; alpha-1-antitrypsin deficiency; Alzheimer's disease; breast cancer; cystic fibrosis; galactosaemia; congenital hypothyroidism; maple syrup urine disease; neurofibromatosis; phenylketonuria; sickle cell disease; Smith-Lemli-Opitz Syndrome; autoimmune disease; shipping fever; cattle; mastitis; bacterial; viral; infection; Types I; type; II; diabetes.

Synthetic.

Location/Qualifiers

Key misc\_feature

1..130 /note= "Remainder of Fl(-) on from pBluescript11 sk(-)"

misc\_feature

131..132 /tag= b /note= "Residue from ligation of restriction enzyme sites used in construction"

promoter

misc\_feature

133..1777 /tag= C /note= "CMV promoter/enhancer from pGM12"

misc\_feature

1780..2987 /note= "CMV promoter is modified by the addition of an ACC sequence upstream of ATG"

misc\_feature

2988..2993 /tag= d /note= "Residue from ligation of restriction enzyme sites used in construction"

misc\_feature

misc\_feature

polyA\_signal

misc\_feature

misc\_feature

misc\_feature

misc\_feature

insertion\_seq

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

/note= "May be used to insert the sequence of interest into the vector"

promoter

4049..5693 /tag= n /note= "CMV promoter/enhancer from pGM12"

misc\_feature

5694..5701 /tag= o /note= "Residue from ligation of restriction enzyme sites used in construction"

CDS

5702..6617 /tag= p /note= "Dated reporter coding sequence including polyA from pDated1.1"

misc\_feature

6618..7101 /tag= q /note= "Multiple cloning site from pBluescript11 sk(-)"

misc\_feature

7102..7106 /note= "May be used to insert the sequence of interest into the vector"

insertion\_seq

7107..7176 /tag= r /note= "Residue from ligation of restriction enzyme sites used in construction"

misc\_feature

7177..7218 /tag= s /note= "Right insertion sequence recognised by Tn10"

misc\_feature

7219..8062 /tag= t /note= "Non-coding lambda DNA residual from pNK2859"

misc\_feature

8063..10263 /tag= u /note= "Non-coding DNA residual from pNK2859"

misc\_feature

10264..10263 /tag= v /note= "Remainder from pBluescript11 sk(-)"

misc\_feature

10264..10263 /tag= w /note= "Remainder from pBluescript11 sk(-)"

misc\_feature

10264..10263 /tag= x /note= "Remainder from pBluescript11 sk(-)"

misc\_feature

10264..10263 /tag= y /note= "Remainder from pBluescript11 sk(-)"

misc\_feature

10264..10263 /tag= z /note= "Remainder from pBluescript11 sk(-)"

misc\_feature

10264..10263 /tag= AA /note= "Remainder from pBluescript11 sk(-)"

misc\_feature

10264..10263 /tag= AB /note= "Remainder from pBluescript11 sk(-)"

misc\_feature

10264..10263 /tag= AC /note= "Remainder from pBluescript11 sk(-)"

misc\_feature

10264..10263 /tag= AD /note= "Remainder from pBluescript11 sk(-)"

misc\_feature

10264..10263 /tag= AE /note= "Remainder from pBluescript11 sk(-)"

Example 2; SEQ ID NO 2; 150pp; English.

This sequence represents a new vector, pTmMod (CMV/Red), comprising a

transposase gene operably linked to a first promoter and one or more gene

of interest operably linked to one or more additional promoters, where

the one or more genes of interest and their operably linked promoters are

flanked by transposase insertion sequences recognized by the transposase

and the first promoter comprises a modified Kozak sequence comprising

ACCATG. The transposase is a Tn10 transposase and is modified in one to

twenty of the first codons. The gene of interest is operably linked to a

second promoter, which is a constitutive promoter or an inducible

promoter, e.g. an ovalbumin or a vitellogenin promoter. The vector

CC further comprises a polyA sequence, a conalbumin polyA sequence, operably  
 CC linked to the transposase gene. The vector further comprises two stop  
 CC codons operably linked to the transposase gene. The first or second gene  
 CC of interest is operably linked to a second or third promoter,  
 CC respectively. The first and a second gene of interest are operably linked  
 CC to a second promoter. The vector further comprises an enhancer operably  
 CC linked to the one or more genes of interest. The enhancer comprises at  
 CC least a portion of an ovalbumin enhancer. The vector further comprises an  
 CC egg directing sequence, ovalbumin or ovomucoid signal sequence or  
 CC vitellogenin targeting sequence, operably linked to the one or more genes  
 CC of interest. The vector is useful in producing transgenic animals with  
 CC desired proteins or molecules. They are also useful in treating  
 CC Huntington's disease, alpha-1-antitrypsin deficiency, Alzheimer's  
 CC disease, breast cancer, cystic fibrosis, galactosemia, congenital  
 CC hypothyroidism, maple syrup urine disease, neurofibromatosis,  
 CC phenylketonuria, sickle cell disease, Smith-Lemli-Opitz Syndrome,  
 CC autoimmune diseases, shipping fever in cattle, mastitis, bacterial or  
 CC viral infections or Types I and II diabetes.

XX Sequence 10263 BP; 2648 A; 2595 C; 2374 G; 2646 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	3,31e-138	Length:	10263
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	12	Gaps:	0

US-10-006-922a-12 (1-225) x AD104102 (1-10263)

QY 1 MetcArgeSerSerlybAenValIlelygGluBheMetArghelybValArghMetGluGly 20

Db 5725 GTGGCGCTCTCCAGAAAGTATCATGAGAGTTCATGCGCTTCAAGTGGCATGAGGCG 5784

QY 21 ThArValAenGlyHISGluPheGluIleGluGlyGluGlyGluGlyGluGlyGluGly 40

Db 5785 ACCGTGAACGGCCAGATTGAGATCCAGGCGCAGGCGCGCGCCCTTACGAGGCG 5844

QY 41 HisAenThVallybLeuLyValThlyrGlyGlyProLeuProPheAlaITPArpIle 60

Db 5845 CACAAACACCGTGAAGCTGAAGGTGACCAAGGCGCGCCCTGCGCTTGGCGGACATC 5904

QY 61 LeuSerProGluPheGluIntyrglySerlybVallyrVallybSHISpRoAlAAspIlePro 80

Db 5905 CTGTCCCGCCAGTTCACATGAGCTCCCAAGGTGACGGAAGCACCCCGCAGACATCCCC 5964

QY 81 AspTyrllyblybLeuSerPheProGluGlyPheLySTPGLuArqValMetAsnPheGlu 100

Db 5965 GACTACAAAGAAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGCTGATGAACTTCAG 6024

QY 101 AepGlyGlyValIValThValThrGlnAepSerSerleuGlnAepGlyCyAspHeIleTy 120

Db 6025 GACGCGCGCGTGTGTACCGTGAACCAAGACCTCTCCGACAGACGCGCTGCTCATCTAC 6084

QY 121 LybVallybPheIleGlyValAsnPheProSerArpGlyProValMetGlnLyblyrThr 140

Db 6085 AAGGTGAAGTTCATCGCGGTAACTTCCCTCCAGCGGCGCGCTGAATGCAAGAAAGACC 6144

QY 141 MetGlyTPGLuAlaSerThrglyblyrlyrProArqAepGlyValIleLybGlyGlu 160

Db 6145 ATGGGCTGGAGGCTTCCACCGAGGCTGTACCCCGCGACGCGCTGTAAAGGCGAG 6204

QY 161 IleHislybAlaLeuLybLeuLybAepGlyGlyHisTyrlleuValGluPheLybSerIle 180

Db 6205 ATCCACAAAGGCCCTGAAGCTGAAGGACGCGCGCCACTACGTGTGAAGTTCAAAGTCCATC 6264

QY 181 TyrlMetAlaLyblyrProValGlnleuProGlyTyrlTyrlTyrlValAAspSerlybLeuAsp 200

Db 6265 TACATGGCGCAAGAGCCGCTGACGTGCGCGCTACTACTACGTGAGATCCAAAGCTGAC 6324

QY 201 IleThrSerHisAenGluAepTyrlThrIleValGluGlnTyrglyblyrThrglyGlyArg 220

Db 6325 ATGACCTCCCAAGACGAGCTACACCATGCTGAGAGTACGAGCGCACCGAGGCGCG 6384

QY 221 HisHisLeuPheLeu 225

Db 6385 CACCACCTGTTCTG 6399

#### RESULT 40

ABR95906

ID ABR95906 standard; cDNA; 681 BP.

AC ABR95906;

DT 29-MAY-2002 (first entry)

DE Yeast optimised RFP encoding cDNA SEQ ID NO 2.

XX Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;

XX Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;

XX Escherichia coli; green fluorescent protein; biotechnology; gene; ss.

XX Anthozoa.

XX Key

XX CDS

XX DE20001395-UI.

XX 15-MAR-2001.

XX 27-JUN-2000; 2000DE-02001395.

XX 27-JUN-2000; 2000DE-02001395.

XX (GPCR-) GPC BIOTECH AG.

XX WPI: 2002-228394/29.

XX P-PSDB; ABB08835.

XX New DNA encoding red fluorescent protein, useful as marker in

XX biotechnology, has sequence optimized for expression in eukaryotes,

XX especially yeast or plants.

XX Claim 2; Fig 1; 19pp; German.

XX The invention relates to DNA (I) containing either sequence ABR95905 or

XX sequence ABR95906 encoding a yeast optimised Red Fluorescent Protein

XX (YRFP). (I) are used to express red fluorescent protein (RFP) in

XX eukaryotes, especially yeast, especially Saccharomyces cerevisiae and

XX plants, especially dicotyledonous plants including Nicotiana tabacum or

XX Arabidopsis thaliana and also in prokaryotes, especially bacteria,

XX especially Escherichia coli. RFP is useful in the same way as green

XX fluorescent protein but is more generally applicable in modern

XX biotechnology. (II) are optimised for expression in yeast and so generate

XX RFP at higher levels with stronger fluorescence and thus lowers the

XX detection limit and gives a better signal-to-noise ratio

XX Sequence 681 BP; 198 A; 147 C; 161 G; 175 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	9,65e-140	Length:	681
Score:	1209.00 <td>Matches:</td> <td>224</td>	Matches:	224
Percent Similarity:	100.00% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	100.00% <td>Mismatches:</td> <td>0</td>	Mismatches:	0
Query Match:	99.59% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	6	Gaps:	0

US-10-006-922a-12 (1-225) x ABR95906 (1-681)

QY 2 ArgSerSerlybAenValIlelygGluBheMetArghelybValArghMetGluGlyThr 21

Db 7 AGATCTTCAAGAACGTCATCAAGAAATCATGAGATTCAAGGTTAAGATGAAAGAGTACT 66

QY 22 ValaAnglyHsAngluPheGlu11Glu1Gly1Glu1Gly1ArgProTyGluGlyHis 41  
 Db 67 GTTAACGGTCACGAATTGAAATCGAAGGTGAAGGTGAAGCACTACGAAGGTAC 126  
 QY 42 AenThrVallybLeuLybValThlybGlyGlyProLeuProPheAlaTrpAspIleLeu 61  
 Db 127 AACACTGTCAGATTGAGATTACTGAAGGGTGTCCATTCGATTCGATTCGATTCGATTCG 186  
 QY 62 SerProGlnPheGlnTyGlySerlybValTyVallybHisProAlaAspIleProAsp 81  
 Db 187 TCTCCACAAATTCCAATACGATTCCTTAAGGTTCTAGCTCAAGCAACCCAGCTGACATTCAGAC 246  
 QY 82 TyrllybLybLeuSerPheProGlnGlyPheLybTrpGluArgValMetAsnPheGluAsp 101  
 Db 247 TACAAGAAAGTTGTCTCTCCAGAAAGTTTCAAGTGGAAAGAGTCATGACATTCGAAAGAC 306  
 QY 102 GlyGlyValValThValThGlnAspSerSerLeuGlnAspGlyCybPheIleTyrls 121  
 Db 307 GGTGGTGTGTTACTGTACTCAAGACTCTCTCTTCAAGACGGTGTTCATCTACACAG 366  
 QY 122 VallybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybSerThMet 141  
 Db 367 GTCAAGTTCATCGGTGTCACTTCCATCTGACGGTCCAGTCATCAAAAGAAAGACTATG 426  
 QY 142 GlyTrpGluAlaSerThrGluArgLeuTyProArgAspGlyValLeuLybGlyGluIle 161  
 Db 427 GGTTCGGAAGCTTCTACCGAAGCTTTGTACCCCAAGAGACGGTGTTCGAAAGGTGAAATC 486  
 QY 162 HisLybAlaLeuLybLeuLybAspGlyGlyHisTyrlLeuValGluPheLybSerIleTyrl 181  
 Db 487 CACAAGGCTTCAAGATTGAAAGACGGTGTCTACTTGTGAAATTCAAAGCTTCTAC 546  
 QY 182 MetAlaLybLybProValGlnLeuProGlyTyTyTyTyValAspSerLybLeuAspIle 201  
 Db 547 ATGGCTTAAGAAAGCCAGTCCAAATTGCCAGGTACTACTACGTTGACTTCAAGTTGACATC 606  
 QY 202 ThrSerHisAngluAspTyThrIleValGluGlnTyGluArgThrgluGlyArgHis 221  
 Db 607 ACCTTTCACAAAGAAAGCTACACTATCGTCGAACATACGAAGCTACGAAGGTAGACAC 666  
 QY 222 HisLeuPheLeu 225  
 Db 667 CACTGTCTCTTG 678  
 Db  
 RESULT 41  
 ABA95907  
 ID ABA95907 standard; RNA; 681 BP.  
 AC ABA95907;  
 XX  
 DT 29-MAY-2002 (first entry)  
 XX  
 DE Yeast optimised RFP encoding RNA SEQ ID NO 3.  
 XX  
 KM Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
 KM Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
 KM Escherichia coli; green fluorescent protein; biotechnology; ss.  
 XX Anthozoa.  
 OS  
 PN DE20001395-U1.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 27-JAN-2000; 2000DE-02001395.  
 XX  
 PR 27-JAN-2000; 2000DE-02001395.  
 XX  
 PA (GPCB-) GPC BIOTECH AG.  
 XX  
 DR WPI, 2002-228394/29.  
 XX

PF New DNA encoding red fluorescent protein, useful as marker in  
 FT biotechnology, has sequence optimized for expression in eukaryotes,  
 PT especially yeast or plants.  
 XX  
 PS Disclosure; Page 11; 19pp; German.  
 XX  
 CC The invention relates to DNA (I) containing either sequence ABA95905 or  
 CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
 CC (YRFP). (I) are used to express red fluorescent protein (RFP) in  
 CC eukaryotes, especially yeast, especially *Saccharomyces cerevisiae* and  
 CC plants, especially dicotyledonous plants including *Nicotiana tabacum* or  
 CC *Arabidopsis thaliana* and also in prokaryotes, especially bacteria,  
 CC especially *Escherichia coli*. RFP is useful in the same way as green  
 CC fluorescent protein but is more generally applicable in modern  
 CC biotechnology. (I) are optimised for expression in yeast and so generate  
 CC RFP at higher levels with stronger fluorescence and thus lowers the  
 CC detection limit and gives a better signal-to-noise ratio. The present  
 CC sequence is that of an RNA sequence corresponding to the yeast optimised  
 CC RFP encoding cDNA given as SEQ ID NO 2 (ABA95906)  
 XX  
 SQ Sequence 681 BP; 198 A; 147 C; 161 G; 0 T; 175 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 9.65e-140 Length: 681  
 Score: 1209.00 Matches: 224  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.59% Indels: 0  
 DB: Gaps: 0  
 US-10-006-922A-12 (1-225) x ABA95907 (1-681)  
 QY 2 ArgSerSerLybAsnValIleLybGluPheMetArgPheLybValArgMetGluGlyThr 21  
 Db 7 AGAUCUUCUAAAGAACGUCUACAGGAUUCAGAUUCAGAUUCAGAUUCAGAUUCAGAUUC 66  
 QY 22 ValaAnglyHsAngluPheGlu11Glu1Gly1Glu1Gly1ArgProTyGluGlyHis 41  
 Db 67 GUUAAAGGUCACGAUUCGAAUUCGAAAGUAGAGUAGAGUAGACUACGAAGGUCAC 126  
 QY 42 AenThrVallybLeuLybValThlybGlyGlyProLeuProPheAlaTrpAspIleLeu 61  
 Db 127 AACACTGTCAGATTGAGATTACTGAAGGGTGTCCATTCGATTCGATTCGATTCGATTCG 186  
 QY 62 SerProGlnPheGlnTyGlySerlybValTyVallybHisProAlaAspIleProAsp 81  
 Db 187 UCUCCACAATTCCAATACGATTCCTTAAGGTTCTAGCTCAAGCAACCCAGCTGACATTCAGAC 246  
 QY 82 TyrllybLybLeuSerPheProGlnGlyPheLybTrpGluArgValMetAsnPheGluAsp 101  
 Db 247 UACAAGAAAGTTGTCTCTCCAGAAAGTTTCAAGTGGAAAGAGTCATGACATTCGAAAGAC 306  
 QY 102 GlyGlyValValThValThGlnAspSerSerLeuGlnAspGlyCybPheIleTyrls 121  
 Db 307 GGTGGTGTGTTACTGTACTCAAGACTCTCTCTTCAAGACGGTGTTCATCTACACAG 366  
 QY 122 VallybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybSerThMet 141  
 Db 367 GTCAGAGUUCACGUGUCACUCCACUCCAGGUCGACGUCACUCCAAAGAAAGACUUVG 426  
 QY 142 GlyTrpGluAlaSerThrGluArgLeuTyProArgAspGlyValLeuLybGlyGluIle 161  
 Db 427 GGUUGGAAAGCUCUACCGAACGUGUUGAACCCAAAGACGUGUGUUGAAGGUGAAUUC 486  
 QY 162 HisLybAlaLeuLybLeuLybAspGlyGlyHisTyrlLeuValGluPheLybSerIleTyrl 181  
 Db 487 CACAAGGCTTCAAGATTGAAAGACGGTGTCTACTTGTGAAATTCAAAGCTTCTAC 546  
 QY 182 MetAlaLybLybProValGlnLeuProGlyTyTyTyTyValAspSerLybLeuAspIle 201  
 Db 547 AUGGCUAAGAAAGCCAGUUCUCCAGGUCUACGUGUACUACGUGUACUACGUGUACUAC 606  
 QY 202 ThrSerHisAngluAspTyThrIleValGluGlnTyGluArgThrgluGlyArgHis 221

DB 607 ACCCCUACAACGAGACUACUACGAGCAUAUACGACUAGAGUAGACAC 666  
QY 222 HistLeupheleu 225  
DB 667 CACUUGUUCUG 678

RESULT 42  
AAD28208  
ID AAD28208 standard; cDNA; 678 BP.  
AC AAD28208;  
XX 22-APR-2002 (first entry)  
DT  
XX Discosoma sp. fluorescent protein ES encoding cDNA.  
DE  
XX Fluorescent timer protein; protein movement; translocation; trafficking;  
KM promoter activity; gene expression; transgenic plant; gene modification;  
KM protein age; ES; ss.  
XX Discosoma sp.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /tag= a  
FT /product= "ES protein"

XX W0200196373-A2.  
PD 20-DEC-2001.  
XX 13-JUN-2001; 2001WO-US019097.  
XX 14-JUN-2000; 2000US-0211607P.  
XX  
XX (CLON-) CLONTECH LAB INC.  
PI Fradkov AF, Terbkikh A;  
XX  
XX WPI; 2002-154595/20.  
DR P-PSDB; AAE17541.  
XX  
PT New fluorescent timer proteins comprising an emission spectrum that  
PT changes over time from a first wavelength to a second wavelength, useful  
PT for monitoring intracellular protein movement, translocation, trafficking  
PT or stability.  
XX  
PS Disclosure; Fig 2; 89pp; English.

XX The invention relates to a fluorescent timer protein having an emission  
XX spectrum that changes over time after synthesis from a first wavelength  
XX to a second wavelength. The fluorescent timer proteins are useful in  
XX monitoring the activity of a promoter, determining the age of a protein,  
XX identifying an agent that modulates the activity of a promoter and in  
XX enriching a population of cells comprising a fluorescent timer protein.  
XX The fluorescent timer proteins are also useful for assessing gene  
XX expression during development of a multicellular organism or during  
XX cellular differentiation, in response to a drug or other inducer of  
XX promoter activity, as a reporter to serve as a read-out of promoter  
XX activity, monitoring intracellular protein movement or translocation,  
XX protein trafficking, or protein stability, to investigate temporal  
XX aspects of the activity of a regulatory element, for determining cell  
XX fate during development and organ remodeling, in spatial and temporal  
XX visualization of newly synthesized proteins and accumulated proteins, and  
XX in distinguishing between newly formed and pre-existing structures, e.g.  
XX membrane junctions and extracellular matrix components. The fluorescent  
XX timer proteins may further be used to investigations where photo bleaching  
XX techniques are employed, as detectable labels, as selectable markers, as  
XX biosensors in prokaryotic and eukaryotic cells, in protease cleavage  
XX assays, and as second messenger detectors. The nucleic acids can be used  
XX to generate transgenic, non-human plants or animals or site-specific gene  
XX modifications in cell lines. The present sequence is a cDNA encoding

CC Discosoma sp. ES fluorescent protein derived from humanised wild-type  
CC Anthozoa protein dfp583 by substituting Val to Ala at 105 and Ser to Thr  
CC at 197  
XX  
SQ Sequence 678 BP; 147 A; 224 C; 203 G; 104 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1,7e-139 Length: 678  
Score: 1207.00 Matches: 223  
Percent Similarity: 99.56% Conservative: 1  
Best Local Similarity: 99.11% Mismatches: 1  
Query Match: 99.42% Indels: 0  
DB: Gaps: 0

US-10-006-922A-12 (1-225) x AAD28208 (1-678)

QY 1 MetArgSerSerLybAnValIleLygIuPhemeArgPheLybValArgMetGluGly 20  
DB 1 ATGGGCTCTCCAAAGAACGTCATCAAGAGTTCATGCGCTTCAAGGCGCGCATGAGAGGC 60  
QY 21 ThrValAenGlyhiEgIuPheGluIleGluGlyGluGlyArgProTyrrGluGly 40  
DB 61 ACCGTGAACGCCACGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTACGAGAGGC 120  
QY 41 HisAsnThrValLybLeuLybValThrLygIyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACACACCGTGAAAGCTGAAGGTCACCAAGGCGCGCCCTGCTTGGCTGGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrrGlySerLybValTyrrValLybHisProAlaAspIlePro 80  
DB 181 CTGTCCCCCGAGTTCGATCGGCTCCAAAGGTGTACGTGAAGCACCCCGCGACATCCCC 240  
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTrpGluValMetLeuPheGlu 100  
DB 241 GACTACAAAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCTGATGAATCTTGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyr 120  
DB 301 GACGCGCGCTGGCGACCGTCGACCCGAGACTCTCTCGAGAGCGCTGCTTCACTTAC 360  
QY 121 LybValLybPheIleGlyValAanPheProSerAspGlyProValMetGlnLybLeuThr 140  
DB 361 AAGGTGAAGTTCACTCGCGCTGAACCTTCCCTCGAGCGCCCGGATGACAGAAAGACC 420  
QY 141 MetGlyTrrPgluAAserThrGluArgLeuTyrProArgAspGlyValIleLybGly 160  
DB 421 ATGGGCTGGAGGCTTCACCGAGCGCTGTACCCCGCAGCGCGCTGTAAGGGCGAG 480  
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180  
DB 481 ATCCACAAAGCCCTGAAGCTGAAGAGACGGCGGCCACTACCTGTGTGAGTTCAGTCCATC 540  
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAanPheLybLeuAsp 200  
DB 541 TACATGGCCAAAGAACCCGTCGAGCTGCCGCTACTACTACGAGACACCAAGCTGGAC 600  
QY 201 IleHisSerHisAenGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCTCCCAACAACGAGACTACCATGTGTGAGCATGACGAGCGCACCGAGGGCGCG 660  
QY 221 HistLeupheleu 225  
DB 661 CACCACTGTTCCTG 675

RESULT 43  
AAH47656  
ID AAH47656 standard; cDNA; 681 BP.  
XX  
AC AAH47656;  
XX  
DT 30-NOV-2001 (first entry)  
XX Anthozoa red fluorescent protein synthetic mutant nucleotide sequence.



XX Fluorescent protein; Anthozoan; fluorescence; marker; FRET; mutant; ss.  
 XX Synthetic.  
 XX MO200162919-A1.  
 XX PD 30-AUG-2001.  
 XX 13-FEB-2001, 2001MO-US004625.  
 XX 23-FEB-2000, 2000US-0184732P.  
 XX (AURO-) AURORA BIOSCIENCES CORP.  
 XX Nelson D, Zamaire E, Tsien R;  
 DR WPI; 2001-557704/62.  
 DR P-PSDB; AAG65510.  
 XX Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise  
 PT functional red fluorescent proteins, and the encoding nucleic acids, with  
 PT key mutations for improving the proteins function.  
 XX Disclosure; Page 86; 90pp; English.  
 XX The invention provides a nucleic acid encoding functional red fluorescent  
 CC protein (II) that differs from the sequence of an Anthozoan red  
 CC fluorescent protein by at least one amino acid substitution, and with  
 CC different fluorescent properties. The red fluorescent protein of the  
 CC invention can be expressed by standard recombinant methodology. (II) are  
 CC used a fluorescent markers and FRET partners. It is used for identifying  
 CC protein-protein interactions. (II) is also suitable for multiplexed  
 CC fluorescent analysis and FRET-based applications using existing Aequorea  
 CC fluorescent proteins. (II) has improved brightness, reduced spectral  
 CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
 CC The key mutations in the encoding nucleic acids provide improved folding,  
 CC brightness, and create (II) with sharper, more defined excitation and  
 CC emission peaks when expressed in mammalian cells. The present sequence  
 CC represents the nucleotide sequence of an improved synthetic mutant of an  
 CC anthozoan fluorescent protein  
 XX  
 XX Sequence 681 BP; 160 A; 201 C; 220 G; 100 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 3 02e-139 Length: 681  
 Score: 1205.00 Matches: 223  
 Percent Similarity: 99.56% Conservative: 1  
 Best Local Similarity: 99.11% Mismatches: 0  
 Query Match: 99.26% Indels: 0  
 Gaps: 0  
 DB: 4  
 US-10-006-922A-12 (1-225) x AAH47656 (1-681)

QY 101 AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120  
 DB 304 GACGCGCGCGTGTGACCGTGAACCAAGACAGCGCTGACGAGCGGCTGTTATCTAC 363  
 QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysPheThr 140  
 DB 364 AAGGTGAAGTTCACTCGCGTGAACCTTCCAGCGACCGCCCCCGTGAAGAGAAAGACC 423  
 QY 141 MetGlyTyrGlnAlaSerThrGlnArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
 DB 424 ATGGCTGGAGGCGCTCCACGAGCGCCTGTACCCCGCGACGCGCTGTGAAAGGGAG 483  
 QY 161 LLeHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGlnPheLysSerIle 180  
 DB 484 ATCCACAAAGCCCTGAAGCTGAAGAGACGGCGCACTACCTGTGGAGTTCAAGTTCATC 543  
 QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
 DB 544 TACATGGCCAAAGAGCCCGTGCAGCTGCGCGCTACTACTACGTGAGCTCCAAAGCTGAC 603  
 QY 201 LLeHisSerHisAsnGlnAspTyrThrIleValGlnGlnTyrGlnArgThrGlnGlyArg 220  
 DB 604 ATCAACGACCAACAGAGACTACACCATCGTGAAGCAGTACGAGAGACCGAGGCGACGG 663  
 QY 221 HisHisLeuPheLeu 225  
 DB 664 CACCACCTGTCTCTG 678  
 RESULT 44  
 ADR30814  
 ID ADR30814 standard; DNA; 7508 BP.  
 XX ADR30814;  
 AC 04-NOV-2004 (first entry)  
 XX  
 DT Zebrafish BDNF gene construct DNA #2.  
 XX  
 DE Brain-derived neurotrophic factor; BDNF; promoter; zebrafish; ds.  
 KW  
 XX Danto rerio.  
 OS Chimeric.  
 OS Unidentified.  
 OS  
 FH Key  
 FT misc\_feature  
 FT 1..20  
 FT /tag= a  
 FT /note= "Vector DNA"  
 FT 15..5104  
 FT /tag= b  
 FT /note= "Expression vector DNA"  
 FT 21..1776  
 FT /tag= c  
 FT /note= "zebrafish BDNF gene 5' flank region"  
 FT 1777..1804  
 FT /tag= d  
 FT /note= "zebrafish BDNF promoter gene"  
 FT 1805..2099  
 FT /tag= e  
 FT /note= "Exon 1c (5' UT) of Zebrafish BDNF gene "  
 FT 2100..2119  
 FT /tag= f  
 FT /note= "Vector DNA"  
 FT 2120..2815  
 FT /tag= g  
 FT /note= "Reporter vector DNA"  
 FT 2816..2820  
 FT /tag= h  
 FT /note= "Linker DNA"  
 FT 2821..5099  
 FT /tag= i  
 FT /note= "3' flank region"



## Alignment Scores:

Pred. No.: 7,86e-137 Length: 6985  
 Score: 1197.00 Matches: 224  
 Percent Similarity: 99.56% Conservative: 1  
 Best Local Similarity: 99.12% Mismatches: 0  
 Query Match: 98.60% Indels: 1  
 Gaps: 0

US-10-006-922a-12 (1-225) x ACA55359 (1-6985)

```

QY 1 MetArgSerSerIysPheAsnValIleIyegIuPheMetArgPheIysValArgMetGluGly 20
DB 2974 GTGGGCTCTCCCAAGAGGATCAAGAGGATTCAGCCCTTCAGAGGTGAGAGGAGG 3033
QY 21 ThrValAenGIyHISgluPheGluIleGIuGIyGIuGIyGIuGIyARgProTyrgIuGIy 40
DB 3034 ACCGTGAACGGCCACGAGATTGAGATTCAGAGGCGAGAGGCGCGCCCTTACGAGAGGC 3093
QY 41 HisAnthrValIyLeuIyValIThrIyegIyGIyProIeuProPheAlaTrpAspIle 60
DB 3094 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGCGCCCTTCGCTCGGAGCATC 3153
QY 61 LeuSerProGInPheGInTyrgIySerIySValTyrrValIySHISProAlaAspIlePro 80
DB 3154 CTGTCCCGCCAGTTCCAGTACGGCTCCAGAGGTGACGTGAAGCACCCGCGACATCCCC 3213
QY 81 AepTyIyIyIyLeuSerPheProGIuGIyPheIySTrPGIuArgValMetAsnPheGIu 100
DB 3214 GACTACAGAGAGCTCTCTTCCCGAGGGCTTCAGGTGGAGCGGCTGATGAACCTTCAG 3273
QY 101 AepGIyGIyValIyValIThrValIThrGInAepSerSerIeuGInAepGIyCyPheIleTy 120
DB 3274 GACGGCGCGCGTGTGACCGGTGACCCAGGACTCTCTCGAGGAGACGCGCTTCATCTAC 3333
QY 121 IySValIySPhelIleGIyValIAsnPheProSerAepGIyProValMetGInIyIyThr 140
DB 3334 AAGGTGAAGTTCAATGGCGGTGAACCTTCCCTCCGAGCGCGCCGTAATGAGAGAACCC 3393
QY 141 MetGIyTrPGIuAlaSerThrGIuArgIeuTyrrProArgAepGIyValIleuIySglu 160
DB 3394 ATGGGCTGGAGGCGCTTCACCGAGGCGCTGTACCCCGCGAGCGGCTGTGAAGGCGAG 3453
QY 161 IleHISlyS-AlaLeuIyLeuIyAepGIyGIyHISrTyIeuValIgluPheIySerI 180
DB 3454 ATCCACAAAGGCGCTGAAGCTGAAGGACGCGCGCATCTCGTGAAGTTCAAGTCCAT 3513
QY 180 eTyIleAlaIyIySProValGInIeuProGIyTyrrTyrrValIaAspSerIyLeuAs 200
DB 3514 CTACATGGCCAGAGAGCCCGTGCAGCTGCCCGCTTACTTACGTGAGACTCCAACTGGA 3573
QY 200 PHeIleSerHISaenGIyAepTyrrThrIleValIgluGInTyrgIuArgThrGIuGIyAR 220
DB 3574 CATCACTCTCCCAAGAGGACTACCATCTGTGAGCAGTACGAGCGACCAAGGCGCG 3633
QY 220 gHISHisIeuPheIeu 225
DB 3634 CCACCACTGTTCTG 3649

```

## RESULT 46

ABLe1142 standard; DNA; 666 BP.

ABLe1142;

18-SEP-2002 (first entry)

Red fluorescent protein reporter DNA fragment.

Alpha-SMA; smooth muscle actin; screening; smooth muscle cell;

myofibroblast; gene therapy; red fluorescent protein; ds.

Synthetic.

XX

```

PN EPI172375-A1.
XX 16-JAN-2002.
XX 22-DEC-2000; 2000EP-00128446.
XX 11-JUL-2000; 2000DE-01033633.
XX 31-OCT-2000; 2000DE-01053879.
XX (ODEN/) ODONTAL M.
XX Odenthal M, Jung D;
XX WPI; 2002-149590/20.
XX
XX New nucleic acid containing regulatory region of the smooth muscle actin
XX gene, useful e.g. for manipulating gene expression in smooth muscle
XX cells.
XX
XX Disclosure; Page 20; 44pp; German.
XX
XX This invention describes a novel nucleic acid (I) comprising: (i) at
XX least one functional region (Ia) from the regulatory region of the alpha-
XX smooth muscle actin (SMA) gene and (ii) at least one additional
XX functional sequence (Ib) operably linked to (Ia). The products of the
XX invention can be used for preparing genetically modified eukaryotic cells
XX or organisms, for isolation and screening of smooth muscle cells,
XX myofibroblasts or related cells, and for manipulation of gene expression
XX and/or cell function in smooth muscle cell or myofibroblasts,
XX particularly for gene therapy. Component (Ia) provides cell-type- or
XX differentiation-specific expression or modulation of genes. This sequence
XX represents a DNA fragment of red fluorescent protein (EGFP) which can be
XX used as a reporter molecule under the control of the alpha-SMA described
XX in the invention
XX
XX Sequence 666 BP; 202 A; 126 C; 176 G; 162 T; 0 U; 0 Other;
SQ

```

## Alignment Scores:

Pred. No.: 3.82e-138 Length: 666  
 Score: 1196.00 Matches: 221  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.52% Indels: 0  
 Gaps: 0

US-10-006-922a-12 (1-225) x ABLe1142 (1-666)

```

QY 5 LysAsnValIleIySgluPheMetArgPheIySValArgMetGluGlyThrValAsnGIy 24
DB 1 AAGAAATGTTATCAAGAGATTCATGAGGTTTAAAGTTCCATGAGAGAACGGTCAATGG 60
QY 25 HisgluPheGluIleGIuGIyGIuGIyArgProTyrgIuGIyHISAnthrVal 44
DB 61 CACGAGTTTGAATATGAAGGAGAGAGAGAGGAGGCGCATACGAAGCCACAAATCCGTA 120
QY 45 IySleuIySValIThrIyegIyGIyProIeuProPheAlaTrpAspIleLeuSerProGIn 64
DB 121 AAGCTTAAGTAAACCAAGGCGGAGCCTTGGCATTTGCTGGGAATTTTGCACACAA 180
QY 65 PheGInTyrgIySerIySValTyrrValIySHISProAlaAspIleProAspTyrrIyS 84
DB 161 TTTCAGTATGAAGCAAGATATATGCAAGCCTTGCAGCATACCGACTATTAANAAG 240
QY 85 LeuSerPheProGIuGIyPheIySTrPGIuArgValMetAsnPheGIuAepGIyGIyVal 104
DB 241 CTGTCAATTTCCAGAGAGATTTAAATGGGAAAGGTCATGAACCTTGAAGACGCTGGCGTC 300
QY 105 ValIThrValIThrGInAepSerSerIeuGInAepGIyCyPheIleTyrrIySValIyS 124
DB 301 GTTACTGTAAACCAAGAGATTCAGTTTGCAGAGTGGCTTTTCATCTACAAAGGTTC 360
QY 125 IleGIyValIaAsnPheProSerAepGIyProValMetGInIyIyThrMetGIyTrPGIu 144

```

```
Db 361 ATTGGCGTGAACCTTCCTCCGATGSACTGTATGCAAAAAGACATGGCGTGGAA 420
Cc 145 A1aserThrg1uaTgLeuTYrPro1a9a9pG1yVal1eulyG1yG1u11eh1e1yA1a 164
Cc 421 GCCAGCACTGAGCTTTGTATCTCCGTGATGGCGCTGTGAAAGAGATTCAATMAAGCT 480
Cc 165 Leu1y1eulya9a9pG1yG1y1h1a1TYrLeu1a1G1uph1e1y9er11eTYrMet1a1a1y 184
Cc 481 CTGAAGCTGAAGAAGCGGTGTCATTACTTACTTGAATTCAAAAGATTATTCATGGCAAG 540
Cc 185 Ly8ProVa1G1n1eupProG1yTYrTYrTYrVa1a9pSer1y1eula9p11eTh9er1h1s 204
Cc 541 AAGCTTGAGCACTACCAAGGGTACTACTATGTGATCCAAACTGATATTAACAAGCCAC 600
Cc 205 Aa9G1ua9pTYrTh1r11e1a1G1uG1nTYrG1uaTYrTh1rG1uG1yA1gH1a1h1e1uPhe 224
Cc 601 AACGAAGCTATACATTCGTTGACAGACAGATGAAAGAACCGAGGAGCCACCATCTGTTCC 660
Cc 225 Leu 225
Cc 661 CTT 663
Db

RESULT 47
ADO78067
ID ADO78067 standard; DNA; 711 BP.
Cc
Cc ADO78067;
Cc
Cc 26-AUG-2004 (first entry)
Cc
Cc Corallimorpharia red fluorescent protein mutant Redi DNA.
Cc
Cc fluorescent protein; red fluorescent protein; green fluorescent protein;
Cc gene expression detection; regeneration; multiple labeling system;
Cc in vivo marker; microinjection assay; taxonomic marker; colour indicator;
Cc food additive; cosmetic; Actinodiscus; Discosoma; gene; ds; mutant; Redi;
Cc coral-like anemone.
Cc
Cc Corallimorpharia.
Cc
Cc Key Location/Qualifiers
Cc CDS 1..711
Cc /product= "Corallimorpharia red fluorescent protein
Cc mutant Redi"
Cc
Cc US2004110225-A1.
Cc
Cc 10-JUN-2004.
Cc
Cc 09-DEC-2002; 2002US-00314936.
Cc
Cc 09-DEC-2002; 2002US-00314936.
Cc
Cc 09-DEC-2002; 2002US-00314936.
Cc
Cc (G1BB/) G1BBS P D L.
Cc (CART/) CARTER R W.
Cc (SCHM/) SCHMALE M C.
Cc
Cc G1bbs PDL, Carter RW, Schmale MC;
Cc
Cc MPI: 2004-467659/44.
Cc
Cc P-RSDB: ADO78068.
Cc
Cc New first mutant fluorescent protein having an optical property
Cc relatively different to a corresponding optical property produced by a
Cc wild-type red or green fluorescent protein, useful for detecting
Cc expression of a gene.
Cc
Cc Claim 9; SEQ ID NO 1; 30pp; English.
Cc
Cc The invention describes a first mutant fluorescent protein (I) comprising
Cc an amino acid sequence that differs from that of a corresponding wild-
Cc type red fluorescent protein by an amino acid substitution, where the
```

Cc first mutation fluorescent protein mutant has an optical property that  
Cc differs relative to the corresponding optical property produced by the  
Cc wild-type red or green fluorescent protein. (I) is useful for detecting  
Cc expression of gene. The polynucleotide (II) encoding (I) is useful for  
Cc detecting expression of a gene which involves introducing (II) into a  
Cc cell or organism, allowing the cell to replicate, and detecting  
Cc expression of the nucleic acid by emission of fluorescent light. The  
Cc expression of the nucleic acid expression is detected in vivo or in  
Cc vitro. (II) is useful for regenerating an animal such as a zebrafish  
Cc which contains (II). (I) is useful as a marker for detecting expression  
Cc of a gene, in biochemical assays, and as reagents. (I) is also useful in  
Cc multiple labeling systems, as in vivo markers such as in mRNA  
Cc microinjection assays, and as taxonomic markers for studies of genetics,  
Cc colour indicators in diagnostic kits, coloured food additives, and  
Cc cosmetics ingredients. (I) is useful in research for up or down  
Cc regulation, to monitor promoter activity, to allow longer term monitoring  
Cc and to localise proteins. (I) has enhanced properties such as  
Cc substantially enhanced fluorescence and reduced toxicity. This sequence  
Cc encodes Redi, a mutant of the red fluorescent protein isolated from an  
Cc aquatic species believed to be either an Actinodiscus or Discosoma  
Cc species.

SQ Sequence 711 BP; 214 A; 131 C; 189 G; 177 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	4,18e-138	Length:	711
Score:	1196.00	Matches:	221
Percent Similarity:	99.11%	Conservative:	2
Best Local Similarity:	98.22%	Mismatches:	2
Query Match:	98.52%	Indels:	0
DB:	12	Gaps:	0

US-10-006-922A-12 (1-225) x ADO78067 (1-711)

```
Cc 1 MetArgSerSer1y9a9a11elyG1uPh1eMecArgPh1e1y9a1a9MecG1uG1y 20
Cc 1 ATGAGTGTTCACAGATGTTATCAAGAGTTTCATGATGATTAAGTTCTGATGAAGGA 60
Cc 21 ThrVa1a9pG1yH1aG1uPh1eG1u11eG1uG1yG1uG1yTYrProTYrG1uG1y 40
Cc 61 ACGTCAATGAGCAAGATTGAAATGAAGAGCAAGAGAGAGAGCCATACGAAGGC 120
Cc 41 H1a9aThrVa11y9a9a11Th1y9a1yG1yPro1eupProPh1a1aTPa9p11e 60
Cc 121 CACATATCCGTAAAGCTTAAGTACCAAGAGGGGACCTTGCCATTGCTGGATATT 180
Cc 61 LeuSerProG1nPh1eG1nTYrG1ySer1y9a1yTYrVa11y9a1a9p11ePro 80
Cc 181 TTGTCAACCAAAATTCAGTATGAGAGCAAGTATATGCAAGCATCCTGCCACATACCA 240
Cc 81 AspTYr1y9a1yLeuSerPheProG1uG1yPh1e1yTYrG1uArgVa1Me1a9pPh1eG1u 100
Cc 241 GACTATAAAGAGCTGATTTCTGAAGGATTTAAATGGAAAGGTCATGAACCTTTGAA 300
Cc 101 AspG1yG1yVa1a1ThrVa11Thrg1na9pSer1eula9pG1yCy9pHe11eTYr 120
Cc 301 GAGGTCGCGCTCGTTACTGTAAACCAAGATTCCAGTTGCAAGATGGCTGTTCACTAC 360
Cc 121 Ly9a1a1y9pHe11eG1yVa1a9pPheProSer1a9pG1yProVa1Me1G1n1y9a1yThr 140
Cc 361 AAGGTCAAGTTCATGTGCGTGAACCTTCTTCGATGSACTGTATTAGCAAAAGAAACA 420
Cc 141 MecG1yTYrG1u1a9eThrg1ua9pLeuTYrPro1a9a9pG1yVal1eulyG1yG1u 160
Cc 421 ATGGGCTGGGAAGCGACCTGAGCGTTGTATCTCGTATAGCGCTGTGAAAGAGAG 480
Cc 161 11eh1e1y9a1a1euly9a9pG1yG1y1h1a1TYrLeu1a1G1uph1e1y9er11e 180
Cc 481 ATTCAATAAGGCTCTGAAGTTGAAGAAGCGTGCATTAACCTGATGAATCAAAACTATT 540
Cc 181 TYrMe1a1a1y9a1y9pProVa1G1n1eupProG1yTYrTYrTYrVa1a9pSer1y9a9p 200
Cc 541 TACATGGCAAAAGAGCTGTGACGCTACAGGGTACTACTATGTGATCCAAACTGGAT 600
```

QY 201 ILeThSerHISaAngluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
 Db 601 ATAAAGAGCCACAAAGACTATACATCGTTGAGCAGTATGAAGAAGACGAGGAGCGC 660  
 QY 221 HLeHISLeuPheLeu 225  
 Db 661 CACCATCTGTTCTT 675

RESULT 48  
 ADO78069  
 ID ADO78069 standard; DNA; 711 BP.  
 AC ADO78069;  
 XX  
 DT 26-AUG-2004 (first entry)  
 XX  
 DE Corallimorpharia red fluorescent protein mutant RedII DNA.  
 XX  
 KW fluorescent protein; red fluorescent protein; green fluorescent protein;  
 KW gene expression detection; regeneration; multiple labeling system;  
 KW in vivo marker; microinjection assay; taxonomic marker; colour indicator;  
 KW food additive; cosmetic; Actinodiscus; Discosoma; gene; ds; mutant;  
 KW RedII; coral-like anemone.  
 XX  
 OS Corallimorpharia.  
 XX  
 PH Key Location/Qualifiers  
 FT 1..711  
 FT /tag= a  
 FT /product= "Corallimorpharia red fluorescent protein  
 FT mutant RedII"  
 XX  
 PN US2004110225-A1.  
 XX  
 PD 10-JUN-2004.  
 XX  
 PF 09-DEC-2002; 2002US-00314936.  
 XX  
 PR 09-DEC-2002; 2002US-00314936.  
 XX  
 PA (GIBB/) GIBBS P D L.  
 PA (CART/) CARTER R W.  
 PA (SCHM/) SCHMALE M C.  
 XX  
 PI GIBBS PDL, Carter RW, Schmale MC;  
 DR WPI, 2004-467659/44.  
 DR P-PSDB; ADO78070.  
 XX  
 PT New first mutant fluorescent protein having an optical property  
 PT relatively different to a corresponding optical property produced by a  
 PT wild-type red or green fluorescent protein, useful for detecting  
 PT expression of a gene.  
 XX  
 PS Claim 10; SEQ ID NO 3; 30bp; English.  
 XX

The invention describes a first mutant fluorescent protein (I) comprising  
 an amino acid sequence that differs from that of a corresponding wild-  
 type red fluorescent protein by an amino acid substitution, where the  
 first mutation fluorescent protein mutant has an optical property that  
 differs relative to the corresponding optical property produced by the  
 wild-type red or green fluorescent protein. (I) is useful for detecting  
 expression of gene. The polynucleotide (II) encoding (I) is useful for  
 detecting expression of a gene which involves introducing (II) into a  
 cell or organism, allowing the cell to replicate, and detecting  
 expression of the nucleic acid by emission of fluorescent light. The  
 expression of the nucleic acid expression is detected in vivo or in  
 vitro. (II) is useful for regenerating an animal such as a zebrafish  
 which contains (II). (I) is useful as a marker for detecting expression  
 of a gene, in biochemical assays, and as reagents. (I) is also useful in  
 multiple labeling systems, as in vivo markers such as in mRNA  
 microinjection assays, and as taxonomic markers for studies of genetics.

CC colour indicators in diagnostic kits, coloured food additives, and  
 CC cosmetics ingredients. (I) is useful in research for up or down  
 CC regulation, to monitor promoter activity, to allow longer term monitoring  
 CC and to localise proteins. (I) has enhanced properties such as  
 CC substantially enhanced fluorescence and reduced toxicity. This sequence  
 CC encodes RedII, a mutant of the red fluorescent protein isolated from an  
 CC aquatic species believed to be either an Actinodiscus or Discosoma  
 CC species.  
 XX  
 SQ Sequence 711 BP; 215 A; 131 C; 188 G; 177 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4,18e-138 Length: 711  
 Score: 1196.00 Matches: 221  
 Percent Similarity: 99.11% Conservative: 2  
 Best Local Similarity: 98.22% Mismatches: 2  
 Query Match: 98.52% Indels: 0  
 DB: 12 Gaps: 0

US-10-006-922a-12 (1-225) x ADO78069 (1-711)

QY 1 MetArgSerSerlySaenValIleLySGluPheMetArgPheLySaValArgMetGluGly 20  
 Db 1 ATGAGTTGTTCCAAAGAAATGTTATCAGAGAGTTCAAGGTTTAAAGTTTCATGCAAGGA 60  
 QY 21 ThrValIleAnglyHISGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
 Db 61 ACGGTCAATGGGACGAGTTGAAATGAAAGCCAAAGAGAAAGAGAGCCATACGAAGC 120  
 QY 41 HisAenThrValLySLeuLySaValThrLySGlyGlyProLeuProPheAlaITrAspIle 60  
 Db 121 CACATACCGTAAAGCTTAAAGTAAACCAAGGGGGAGACCTTTGCCATTGCTTGGGATATT 180  
 QY 61 LeuSerProGlnPheGlnTyrGlySerLySaValTyrValLySHISProAlaAspIlePro 80  
 Db 181 TTGTACACCAATTTCAATGATGAAGCAAGGATATGCAAGCATCCCGCAATACCA 240  
 QY 81 AspTyrLyLySLeuSerPheProGluGlyPheLySTrPGluArgValMetAspPheGlu 100  
 Db 241 GACTATMAAAGCGTCATTTCTGAAAGATTTAAATGAAAGGCTCATGAACTTTGAA 300  
 QY 101 AspGlyGlyValIleValThrValThrGlnAspSerLeuGlnAspGlyCySPheIleTyr 120  
 Db 301 GACGCTGCGCTGCTTACGTAAACCAAGATTCAGATTGCGAGATGGCTGTTCACTTAC 360  
 QY 121 LySValLySPhelIleGlyValAenPheProSerAspGlyProValMetGlnLySlyThr 140  
 Db 361 AAGTCAAGTTCATTTGGCGTGAACTTCTCTGATGACCTGTTATGCNAAGAAAGACA 420  
 QY 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuLySGlyGlu 160  
 Db 421 ATGGGCTGGGAAGCCAGCACTGAGGTTTGTACTTCGTAAGGGGCTGTTGAAAGAGAG 480  
 QY 161 ILeHISLySaLeuLySLeuLySaAspGlyGlyHISLyTyrLeuValGluPheLySerIle 180  
 Db 481 ATTCAITAAAGCTCTGAAGTTGAAAGACGCTGCTATTACCTTGAATGAAATCAAACTATT 540  
 QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrValAspSerLySLeuAsp 200  
 Db 541 TACATGGCAAAAGAACCTGTGACGTACCAAGGGTACTTATAGTTGATGACCCAAACGTGAT 600  
 QY 201 ILeThSerHISaAngluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
 Db 601 ATAAAGAGCCACAAAGACTATACATCGTTGAGCAGTATGAAGAAGACGAGGAGCGC 660  
 QY 221 HLeHISLeuPheLeu 225  
 Db 661 CACCATCTGTTCTT 675

RESULT 49  
 AD284222  
 ID AD284222 standard; DNA; 711 BP.  
 XX

AC	AD284222;
XX	14-JUL-2005 (first entry)
XX	Actinodiscus/Discosoma RFP mutant Red I-encoding DNA, SEQ ID NO:1.
DE	Red fluorescent protein; Red I; mutant; gene; de.
XX	CoralImmunopharia.
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	1..711
FT	/tag= a
FT	/product= "Mutant Ac/DsRFP Red I"
FT	replace(694,A)
FT	/tag= b
FT	/note= "this nucleotide is A in the Red II mutant,
FT	resulting in Aan rather than Asp at position 232 of the
FT	protein"
XX	
PN	US2005100954-A1.
PD	
PD	12-MAY-2005.
XX	
XX	23-DEC-2004; 2004US-00021014.
PP	
PR	09-DEC-2002; 2002US-00314936.
XX	
PA	(GIBB/) GIBBS P D L.
PA	(CARTER) CARTER R W.
PA	(SCHM/) SCHMALE M C.
PI	Gibbs PDL, Carter RW, Schmale MC;
DR	WPI; 2005-345399/35.
DR	P-PDB; AD284223.
PT	
PT	Novel mutant fluorescent protein comprising amino acid sequence that
PT	differs from wild-type red or green fluorescent protein by amino acid
PT	substitutions, useful as markers for detecting desired gene expression.
XX	
PS	Claim 3; SEQ ID NO 1; 29pp; English.
XX	
XX	The invention relates to mutant red and green fluorescent proteins which
CC	have higher fluorescence intensities compared to the wild-type proteins
CC	from which they are derived. The invention also discloses nucleic acids
CC	encoding the mutant fluorescent proteins, and transgenic animals
CC	comprising such a nucleic acid. The mutant red fluorescent proteins (RFP)
CC	of the invention, Red I (AD284223) and Red II (AD284225), are derived
CC	from a novel wild-type RFP (referred to as Ac/DsRFP) isolated from a
CC	mushroom coral believed to be either an Actinodiscus or Discosoma
CC	species, while the mutant green fluorescent proteins (GFP), Green I
CC	(AD284227) and Green (AD284229), are derived from a novel wild-type GFP
CC	(known as McGFP) isolated from the great star coral Montastraea
CC	cavernosa. The Ac/DsRFP and McGFP mutants were obtained by subjecting the
CC	wild-type Ac/DsRFP and McGFP cDNAs to low-stringency PCR to introduce
CC	random mutations, cloning the PCR products into bacterial expression
CC	vectors, and selecting the transformed bacteria having higher relative
CC	fluorescence compared to those expressing the wild-type proteins. The
CC	mutant fluorescent proteins of the invention may be used in a wide range
CC	of applications. They are useful as markers for detecting the expression
CC	of a gene of interest, or can be used as in vivo markers in mRNA
CC	microinjection assays in transgenic animals. They can be used in
CC	expression systems (e.g., as components of fusion proteins), in multiple
CC	labeling systems, or can be used in biochemical assays and as reagents.
CC	They may additionally be used as taxonomic markers for studies of
CC	candarian genetics, as color indicators in diagnostic kits, as colored
CC	food additives and as cosmetic ingredients. The present sequence
CC	represents DNA encoding the mutant Actinodiscus/Discosoma red fluorescent
CC	protein Red I.
XX	
SQ	Sequence 711 BP; 214 A; 131 C; 189 G; 177 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	4,18e-138	Length:	711
Score:	1196.00	Matches:	221
Percent Similarity:	99.11%	Conservative:	2
Best Local Similarity:	98.22%	Mismatches:	2
Query Match:	98.52%	Gaps:	0
DB:	14	Indels:	0

  

US-10-006-922A-12 (1-225) x ADZ84222 (1-711)	
QY	1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluLY 20
Dd	1 ATGAGTGTTCACAAAGATGTATTCAAGAAGTTCATGAAGTTTAAGCTTCATGGAAACA 60
OY	21 ThrValAsnGLYHileGluPheGluIleGluGlyGluGlyGluGlyArgProTYrGluGly 40
Dd	61 ACGGTCAATGGGCACAGTTTGAATAAGAACGCCAAGAAAGGAGGCCATACCAGGC 120
OY	41 HisAsnThrValLysLeuLysValThrlYsgLygLYProLeuProPheLatrapApille 60
Dd	121 CACAAATRCCTTAAGCTTTAAGSTAACAAAGGGGGAACCTTGCCATTGGCTTGGATATT 180
OY	61 LeuSerProGlnPheGlnTYrgLserLYsVALTYrValLysHisProAlasppIlePro 80
Dd	181 TTGTCAACCAAAATTTCAATGATGAAGCAAGATATGTCAAGACCTCGCAGCATACCA 240
OY	81 AspTYrLYsLYsLeuSerPheProGluGlyPheLYsTPRGLuARGValMetAsnPheGlu 100
Dd	241 GACTATATAAAAGCTGCATTTCTCGTAGAGATTTAATGGAAAGGGTCATGAACTTTAA 300
OY	101 AspGLYGLYValValThrValThrgInAspSerSerLeuGlnAspGLYCysPheIleTYr 120
Dd	301 GACGGTGGCCTCTTACTTACTGAACCAAGATTCACATTTGACAGAAAGGCTGTTTCACTAC 360
OY	121 LysValLysPheIleGlyValAsnPheProSerAspGLYProValMetGlnLysLYsThr 140
Dd	361 AAGGTCAAAGTTCATGGCGCTGAACCTTTCTTCATGACCTGTTATCCAAGAAAGACA 420
OY	141 MetGLYTRGInLysSerThrgInuArgLeuTYrProArgAspGLYValLysLYsGlu 160
Dd	421 ATGGCGTGGGAACCGACACTGAGCCTTGATCTCGTATGAGCGCTGTGAAGAGAG 480
OY	161 IleHisLYsValLeuLysLeuLysAspGLYGLYHISLYrLeuValGluPheLysSerIle 180
Dd	481 ATTCAATAGGCTCTGAAGTTGAAGAAGCGGTGCATTACCTAGTGAATTCAAAACTATT 540
OY	181 TyrMetAlaLysLYsProValGlnLeuProGLYTYrTYrTYrValAspSerLYsLeuAsp 200
Dd	541 TACATGGCAAGAAAGCCTGTGCAGCTACACAGGGGTCTACTAATGTGACTCCAAATGGAT 600
OY	201 IlleHserHisAsnGluAspTYrThrIleValGluGlnTYrGluArgThrgLugLYArg 220
Dd	601 ATAACAAGCCACAACAAAGACTATACATCGTTGACAGCTATGAAGAAGACGAGGAGCC 660
OY	221 HisHisLeuPheLeu 225
Dd	661 CACCATCTGTTCCCTT 675

  

RESULT 50	
ADZ84224	
ID	ADZ84224 standard; DNA; 711 BP.
XX	
AC	ADZ84224;
DT	14-JUL-2005 (first entry)
XX	
DE	Actinodiscus/Discosoma RFP mutant Red II-encoding DNA, SEQ ID NO:3.
XX	
KW	Red fluorescent protein; Red II; mutant; gene; ds.
XX	
OS	CoralJimopharia.
	Synthetic.

XX Key Location/Qualifiers  
 FH 1..711  
 FT /\*tag= a  
 FT /product= "Mutant Ac/DsRFP Red II"  
 FT replace(694,G)  
 FT /\*tag= b  
 FT /note= "This nucleotide is G in the Red I mutant,  
 FT resulting in Asp rather than Asn at position 232 of the  
 FT protein"  
 XX US2005100954-A1.  
 XX PD 12-MAY-2005.  
 XX PE 23-DEC-2004; 2004US-00021014.  
 XX PR 09-DEC-2002; 2002US-00314936.  
 XX PA (GIBB/) GIBBS P D L.  
 XX PA (CART/) CARTER R W.  
 XX PA (SCHM/) SCHMALE M C.  
 XX GIbbB PDL, Carter RW, Schmale MC;  
 XX MPI; 2005-345399/35.  
 XX DR P-PSDB; AD284425.  
 XX PT Novel mutant fluorescent protein comprising amino acid sequence that  
 XX differs from wild-type red or green fluorescent protein by amino acid  
 XX substitutions; useful as markers for detecting desired gene expression.  
 XX PS Example 4; SEQ ID NO 3; 29pp; English.

XX The invention relates to mutant red and green fluorescent proteins which  
 CC have higher fluorescent intensities compared to the wild-type proteins  
 CC from which they are derived. The invention also discloses nucleic acids  
 CC encoding the mutant fluorescent proteins, and transgenic animals  
 CC comprising such a nucleic acid. The mutant red fluorescent proteins (RFP)  
 CC of the invention, Red I (AD284423) and Red II (AD284425), are derived  
 CC from a novel wild-type RFP (referred to as Ac/DsRFP) isolated from a  
 CC mushroom coral believed to be either an *Actinodiscus* or *Discosoma*  
 CC species, while the mutant green fluorescent proteins (GFP), Green I  
 CC (AD284427) and Green (AD284429), are derived from a novel wild-type GFP  
 CC (known as MGFP) isolated from the great star coral *Montastraea*  
 CC *cavernosa*. The Ac/DsRFP and MGFP mutants were obtained by introducing the  
 CC wild-type Ac/DsRFP and MGFP cDNAs to low-stringency PCR to introduce  
 CC random mutations, cloning the PCR products into bacterial expression  
 CC vectors, and selecting the transformed bacteria having higher relative  
 CC fluorescence compared to those expressing the wild-type proteins. The  
 CC mutant fluorescent proteins of the invention may be used in a wide range  
 CC of applications. They are useful as markers for detecting the expression  
 CC of a gene of interest, or can be used as in vivo markers in mRNA  
 CC microinjection assays in transgenic animals. They can be used in  
 CC expression systems (e.g., as components of fusion proteins), in multiple  
 CC labeling systems, or can be used in biochemical assays and as reagents.  
 CC They may additionally be used as taxonomic markers for studies of  
 CC eukaryotic genetics, as color indicators in diagnostic kits, as colored  
 CC food additives and as cosmetic ingredients. The present sequence  
 CC represents DNA encoding the mutant *Actinodiscus*/*Discosoma* red fluorescent  
 CC protein Red II. Red II was found to have a fluorescent intensity at least  
 CC 50% greater than the mutant RFP Red I (AD284423).

XX Sequence 711 BP, 215 A; 131 C; 188 G; 177 T; 0 U; 0 Other;

# Alignment Scores:

XX Pred. No.: 4,18e-138 Length: 711  
 XX Score: 1196.00 Matches: 221  
 XX Percent Similarity: 99.11% Conservative: 2  
 XX Best Local Similarity: 98.22% Mismatches: 2  
 XX Query Match: 98.52% Indels: 0  
 XX DB: 14 Gaps: 0

US-10-006-922a-12 (1-225) x AD284424 (1-711)  
 QY 1 MetArgSerSerIysAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20  
 DB 1 ATGAGTTGTTCCAAAGAAATGTTATCAAGAGTTCAATGAGTTTAAAGTTGATGAAAGGA 60  
 QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyValGluGlyGluGlyArgProTyrGluGly 40  
 DB 61 ACGGTCAATGGGCACGAGTTTGAAATGAAAGCCAAAGGAGAAAGGAGGCCCTACGAAGGC 120  
 QY 41 HisAsnThrValIysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
 DB 121 CACAAATCCGTAAGCTTAAGTTAACCAAGGGGAGACCTTGGCATTTGCTTGGAAATTT 180  
 QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro 80  
 DB 181 TTGTACACCAATTTCAATGATGAAAGCAAGGATATATGCAAGCATCTCGCCACATACCA 240  
 QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
 DB 241 GACTATMAAAAGCTGTCATTTCTGAAAGATTAAATGGAAAGGTCATGAACTTTGAA 300  
 QY 101 AspGlyGlyValIleThrValThrGlnAspSerLeuGlnAspGlyCysPheIleTyr 120  
 DB 301 GACGTGGCGTCCTTACTGTAACCAAGATTCAGTTTGACGATGGCTTTTCACTTAC 360  
 QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr 140  
 DB 361 AAGTCAAAGTTCAATGGCGTGAACCTTCTCTGATGACCTGTTATGCCAAAGAACCA 420  
 QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleLysGlyLys 160  
 DB 421 ATGGCTGGGAAAGCCAGCATGAGCGTTGTATCTCGTAATGGCGTTTAAAGAGAG 480  
 QY 161 IleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGluPheLysSerIle 180  
 DB 481 ATTCAATAGGCTCTGAATTTGAAGACCGTGTCTATTCCTTGAATTCAAACCTAAT 540  
 QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
 DB 541 TACATGGCAAGAAAGCTGTACAGCTACAGGGTACTCTATGTGTGATCCAAACTGAT 600  
 QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
 DB 601 ATTAACAAACCCCAACAAAGACTATACATCGTTGACGATGAAAGAACCGAGGACGC 660  
 QY 221 HisHisLeuPheLeu 225  
 DB 661 CACCATCTGTTCTT 675  
 RESULT 51  
 AAD46282  
 ID AAD46282 standard; DNA; 678 BP.  
 XX AAD46282;  
 XX 27-DEC-2002 (first entry)  
 XX DE *Discosoma* sp. drfp583 (NFP-6) mutant DNA, ES-NA.  
 XX KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
 KW biosensor; fluorescence activated cell sorting application; fluorescent timer;  
 KW colouring agent; recombinant DNA application; analyte detection assay;  
 KW sunscreen; second messenger detector; drfp583 protein; NFP-6; gene;  
 KW mutant; ds.  
 XX KW *Discosoma* sp.  
 XX OS Synthetic.  
 XX XX Location/Qualifiers  
 FH Key 1..678  
 FT CDS /\*tag= a

PT /product= "drfp583 mutant protein, B5-NA"  
 XX WO200268459-A2.  
 FN 06-SEP-2002.  
 PD 20-FEB-2002; 2002WO-US005749.  
 XX 21-FEB-2001; 2001US-0270983P.  
 PR 04-DEC-2001; 2001US-00006922.  
 XX (CLON-) CLONTECH LAB INC.  
 PA  
 PI Lukyanov S, Lukyanov K, Yamsheovich Y, Savitsky A, Fradkov A;  
 XX WPI; 2002-691654/74.  
 DR P-PSDB; AAE28837.  
 XX  
 PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
 PT of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
 PT analyte detection assays or fluorescence activated cell sorting  
 PT applications.  
 XX  
 PS Disclosure; Page 73; 80pp; English.  
 XX  
 CC The invention relates to nucleic acid molecules encoding non-aggregating  
 CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
 CC useful in analyte detection assays, as colouring agents, as markers in  
 CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, in screening assays, as second  
 CC messenger detectors, in fluorescence activated cell sorting applications,  
 CC in protease cleavage assays or as fluorescent timers. The present  
 CC sequence is *Discosoma* sp. drfp583 (NFP-6) mutant DNA of the invention  
 XX  
 SQ Sequence 678 BP; 145 A; 225 C; 203 G; 105 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 1.63e-137 Length: 678  
 Score: 1191.00 Matches: 220  
 Percent Similarity: 98.67% Conservative: 2  
 Best Local Similarity: 97.78% Mismatches: 3  
 Query Match: 98.11% Indels: 0  
 DB: Gaps: 0  
 US-10-006-922a-12 (1-225) x AAD46282 (1-678)  
 QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20  
 DB 1 ATGGCCTCTCCGAGAACTGATCATCGAGTTTCATGCGCTTCAAGGTGCGATGAGGCG 60  
 QY 21 ThrValaAngIyHsGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40  
 DB 61 ACCGTGAAGCGGCACGAGTTGAGATCGAGGCGCGAGGCGCGCTTACGAGGCGC 120  
 QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
 DB 121 CACAACACCGGAGGTTAAGGTGACCAAGGCGCGCGCTTCCCTTGGCTGGAGACATC 180  
 QY 61 LeuSerProGluPheGluIleTyrglySerIysValTyrglyValIysHsIleProAlaAspIlePro 80  
 DB 181 CTGTCCTCCCGAGTTCACATGCGCTCCCAAGGTGACGTAAGGAGCAGCCGCGCATCCCC 240  
 QY 81 AspTyrlsIysIysLeuSerPheProGluGlyPheIysTrpGluIleValMetAsnPheGlu 100  
 DB 241 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATCTTCAG 300  
 QY 101 AspGlyGlyValIleValThrValThrGlnAspSerSerIleuGlnAspGlyCysPheIleIyr 120  
 DB 301 GACGCGCGCGGCGGACCGGTGACCCAGGACTCTTCCCTGCGAGGACGGCTTCATCTAC 360  
 QY 121 IysValIysPheIleGlyValIAsnPheProSerAspGlyProValMetGlnIysIysThr 140

DB 361 AAGGTAAAGTTTCATCGGCGGTGAACCTTCCCTCCGACGCGCCCGGTAGCAGAAAGACC 420  
 QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrlProArgAspGlyValIleuIysGlyGlu 160  
 DB 421 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCCGCGACGCGTGAAGGGGCGAG 480  
 QY 161 IleHisIysAlaLeuIysLeuIysAspGlyGlyHisIleTyrlleuValGluPheIysSerIle 180  
 DB 481 ATCCACAAAGCCCTGAAGCTGAAGAGAGCGGCGCATCACTGGTGAAGTTCAAGTCATC 540  
 QY 181 TyrMetAlaIysIysPheProValGluLeuProGlyTyrlTyrlValAspSerIysLeuAsp 200  
 DB 541 TACATGGCCAAAGAGCCCGGTGACGTGCCCGGCTACTACTACGTGACACCAAGCTGAC 600  
 QY 201 IleThrSerHisAsnGluuAspTyrlThrIleValGluGlnIleTyrgluArgThrGluIyArg 220  
 DB 601 ATCACTCCACCAACGAGACTACACCATGTGTGAGAGATACAGACCGACCGAGGCGCG 660  
 QY 221 HisIleuPheLeu 225  
 DB 661 CACCACTGTCTCTG 675  
 RESULT 52  
 AAD28209  
 ID AAD28209 standard; cDNA; 678 BP.  
 XX  
 AC AAD28209;  
 XX  
 DT 22-APR-2002 (first entry)  
 XX  
 DE *Discosoma* sp. non-aggregating mutant timer protein B5NA encoding cDNA.  
 XX  
 KW Fluorescent timer protein; protein movement; translocation; trafficking;  
 KW promoter activity; gene expression; transgenic plant; gene modification;  
 KW protein age; mutant; B5NA; ss.  
 XX  
 OS *Discosoma* sp.  
 OS Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT 1..678  
 FT CDS /\*tag= a  
 FT /product= "Mutant B5NA protein"  
 XX  
 FN WO200196373-A2.  
 XX  
 PD 20-DEC-2001.  
 XX  
 PF 13-JUN-2001; 2001WO-US019097.  
 XX  
 PR 14-JUN-2000; 2000US-0211607P.  
 XX  
 PI (CLON-) CLONTECH LAB INC.  
 PA  
 PI Fradkov AF, Tersikh A;  
 XX  
 DR WPI; 2002-154595/20.  
 DR P-PSDB; AAE17542.  
 XX  
 PT New fluorescent timer proteins comprising an emission spectrum that  
 PT changes over time from a first wavelength to a second wavelength, useful  
 PT for monitoring intracellular protein movement, translocation, trafficking  
 PT or stability.  
 XX  
 PS Disclosure; Fig 3; 89pp; English.  
 XX  
 CC The invention relates to a fluorescent timer protein having an emission  
 CC spectrum that changes over time after synthesis from a first wavelength  
 CC to a second wavelength. The fluorescent timer proteins are useful in  
 CC monitoring the activity of a promoter, determining the age of a protein,  
 CC identifying an agent that modulates the activity of a promoter and in  
 CC enriching a population of cells comprising a fluorescent timer protein.  
 CC The fluorescent timer proteins are also useful for assessing gene



expression during development of a multicellular organism or during cellular differentiation, in response to a drug or other inducer of promoter activity, as a reporter to serve as a read-out of promoter activity, monitoring intracellular protein movement or translocation, protein trafficking, or protein stability, to investigate temporal aspects of the activity of a regulatory element, for determining cell fate during development and organ remodeling, in spatial and temporal visualisation of newly synthesised proteins and accumulated proteins, and in distinguishing between newly formed and pre-existing structures, e.g. membrane junctions and extracellular matrix components. The fluorescent timer proteins may further be used to investigate components where photo bleaching techniques are employed, as detectable labels, as selectable markers, as biosensors in prokaryotic and eukaryotic cells, in protease cleavage assays, and as second messenger detectors. The nucleic acids can be used to generate transgenic, non-human plants or animals or site-specific gene modifications in cell lines. The present sequence is a cDNA encoding *Discooma* sp. B5NA non-aggregating fluorescent timer protein derived from fluorescent timer protein B5 by substituting amino acids at positions R2A, K5E and K9T. E5 protein is derived from humanised wild-type *Anthonzoa* protein drFP583 by substituting Val to Ala at 105 and Ser to Thr at 197

## Alignment Scores:

Pred. No.: 1,63e-137 Length: 678  
Score: 1191.00 Matches: 220  
Percent Similarity: 98.67% Conservative: 2  
Best Local Similarity: 97.78% Mismatches: 3  
Query Match: 98.11% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AAD28209 (1-678)

```

QY      1 MetArgSerSerLeuValIleLeuGluPheMetArgPheValArgMetGluGly 20
DB      1 ATGCGCTCTCCGAGAACGTCATCCAGATTCAAGCGCTTAAAGTGGCATGAGAGGCC 60
QY      21 ThrValaAngLYHIGluPheGluIleGluGluGluGluGluGluGluGluGluGluGlu 40
DB      61 ACCGGAACGGCCAGAGTTCAGATCCAGAGCGGCGAGGCGGCGGCGGCGGCGGCGGCGG 120
QY      41 HIsaenThrValLeuLeuValThrLYGlyGlyProLeuProPheAlaTrpAspIle 60
DB      121 CACAACACCGTGAAGCTGAACAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY      61 LeuSerProGlnPheGlnTrpGlySerLeuValTrpValIleHisProAlaAspIlePro 80
DB      181 CTGTCCCCCAAGTTCAGATCCAGCTCCAAAGTGAAGTGAAGACCCCGCGCAATCCCC 240
QY      81 AspTYrLYLeuLeuSerPheProGluGlyPheLYSTrGluArgValMetAsnPheGlu 100
DB      241 GACTTCAGAAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGACGGGTATGAATTCGAG 300
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyAspHeileTYr 120
DB      301 GACGGCGGGGCGGCGAGCCGTGAACCGAGACTCTCCCTGCAAGACGGCTTCAATCTAC 360
QY      121 LyuValLYPheHeileGlyValaenPheProSerAspGlyProValMetGlnLYLeuSThr 140
DB      361 AAGGGAAGATTCAATCGCGTGAATCTTCCCTCGAGAGGCGCGCGTGAATGCAAGAAAGACC 420
QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTYrProArgAspGlyValLeuLYGlyGlu 160
DB      421 ATGGGCTGGAGGCGCTCCACCAAGCGGCTGACCCCGGAGCGGCGGCTGTGAAGGGCGAG 480
QY      161 ILeuLYLeuAlaLeuLYLeuLYAspGlyGlyHisTYrLeuValGluPheLYSerIle 180
DB      481 ATCCACAAAGCGCTGAAGCTGAAGAGCGGCGGCGCACTACCTGGTGAAGTTCAAGTCAATC 540
QY      181 TYrMetAlLYLeuLYProValGlnLeuProGlyTYrTYrTYrValaAspSerLYLeuAsp 200
DB      541 TACATGGCCAAAGAACCCGTCAGCTGCCCGGCTACTACTACTGAGCAACAAAGCTGAGAC 600

```

```

QY      201 ILeuSerHisAngLYuAspTYrThrIleValGluGlnTYrGluArgTrpGluGlyArg 220
DB      601 ATCACTCCCAACAACGAGACTACACCATGTGGAGCATGACAGGCCACCGAGGCCGCC 660
QY      221 HisIleuPheLeu 225
DB      661 CACCACTGTTCCTCG 675

```

## RESULT 53

AAD46281  
ID AAD46281 standard; DNA; 675 BP.

AC AAD46281;

DT 27-DEC-2002 (first entry)

DE *Discooma* sp. drFP583 (NFP-6) mutant DNA, E57-NA.

KM Fluorescent protein; chromoprotein; protease cleavage assay; filter;

KM fluorescence activated cell sorting application; fluorescent timer;

KM biosensor; fluorescence resonance energy transfer application; FRRET;

KM colouring agent; recombinant DNA application; analyte detection assay;

KM sunsreen; second messenger detector; drFP583 protein; NFP-6; gene;

KM mutant; ds.

OS *Discooma* sp.

OS Synthetic.

XX Key

XX CDS

XX WO200268459-A2.

XX 06-SEP-2002.

XX 20-FEB-2002; 2002WO-US005749.

XX 21-FEB-2001; 2001US-0270983P.

XX 04-DEC-2001; 2001US-00006922.

XX (CLON-) CLONTECH LAB INC.

XX Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Pradkov A;

XX WPI, 2002-691654/74.

XX P-PSDB; AAE28836.

XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant

XX of an aggregating Cnidarian chromo- or fluorescent protein or mutant for

XX applications.

XX Disclosure; Page 73; 80pp; English.

XX The invention relates to nucleic acid molecules encoding non-aggregating

XX chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are

XX useful in analyte detection assays, as colouring agents, as markers in

XX recombinant DNA applications, as sunscreens or filters, in fluorescence

XX resonance energy transfer (FRET) applications, as biosensors in

XX prokaryotic and eukaryotic cells, in screening assays, as second

XX messenger detectors, in fluorescence activated cell sorting applications,

XX in protease cleavage assays or as fluorescent timers. The present

XX sequence is *Discooma* sp. drFP583 (NFP-6) mutant DNA of the invention

XX SQ Sequence 675 BP; 142 A; 227 C; 204 G; 102 T; 0 U; 0 Other;

XX Alignment Scores: Pred. No.: 6.74e-137 Length: 675

XX Score: 1186.00 Matches: 219

Percent Similarity: 98.22% Conservative: 2  
 Best Local Similarity: 97.33% Mismatches: 4  
 Query Match: 97.69% Indels: 0  
 DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AAD46281 (1-675)

```

QY 1 MetArgSerSerLybAnVal11leLybGluPhenMetArgPheLybValArgMetGluGly 20
DB 1 ATGGCTCTCTCCGAAAGCATCATCACGAGTTCAATGCGCTTCAAGGTCGATGAGGCG 60
QY 21 ThrValAsnGlyH1sgLupHegLui1egLugLyGluGlyArgProTyrgLugLy 40
DB 61 ACCGTGAACGGCCACGAGTTGAGATCGAGGCCAGGGCCGAGGCCCGCCCTTACAGAGGC 120
QY 41 HisAnThrValLybLeuLybVal1ThrLybGlyGlyProLeuProPheAlaTrpAsp1le 60
DB 121 CACAAACACCGTGAAGCTGAAAGTGAACCAAGGGGGGCGCCCTGCTTCGCTGGAGACATC 180
QY 61 LeuSerProGlnPheGlnTyrgLySerLybValTyrgValLybHisProAlaAsp1lePro 80
DB 181 CTGTCCCCCGCCAGTTCAGATGAGCTCCAAAGGTGAAGTGAACCCCGCGACATCCCC 240
QY 81 AspTyrgLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100
DB 241 GACTACAAAGAACTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTATGAATCTTGAG 300
QY 101 AspGlyGlyValVal1ThrVal1ThrGlnAspSerSerLeuGlnAspGlyCyaspHe1leTy 120
DB 301 GACGCGCGCGTGGCGACCGTGAACCAAGGACTCTCCCTCGCAGAGACGGCTGCTCATCTAC 360
QY 121 LybValLybPhe1leGlyVal1AsnPheProSerAspGlyProValMetGlnLybLybThr 140
DB 361 AAGGTGAAGTTCATCGGGGTAACTTCCCTCCAGCGGCCCGTGAATGCAAAAGACC 420
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrgProArgAspGlyVal1LeuLybGlyGlu 160
DB 421 ATGGGCTGGAGGGCTCCACCGAGCGCTGTACCCTCCCGCAGCGCGCTGTAAAGGCGAG 480
QY 161 1leHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrgLeuVal1GluPheLybSer1le 180
DB 481 ACCCAACAAAGCCCTGAAGCTGAAGGACGGCGGCCACTACGTGTGAGTTCAAGTCCATC 540
QY 181 TyrgMetAlaLybLybProVal1GlnLeuProGlyTyrgTyrgTyrgVal1AspSerLybLeuAsp 200
DB 541 TACATGGCCAAAGCCCGTGCAGCTGCCGCTACTACTACGTGAGCGCAAGCTGGAC 600
QY 201 1leThrSerHisAsnGluAspTyrgThr1leVal1GluGlnTyrgLugLybArgThrgLugLy 220
DB 601 ATCACTCTCCCAACAGAGACTACACCATCTGTGAGCACTACGAGCGCACCGAGGGCGCG 660
QY 221 HisHisLeuPheLeu 225
DB 661 CACCACTGTTCTTG 675

```

# RESULT 54

ADV96699 ID ADV96699 standard; DNA; 747 BP.

AC ADV96699;

XX 10-MAR-2005 (first entry)

XX Plant cell gene silencing related dased gene coding region.

XX gene silencing; plant; short interfering RNA; siRNA; RNA interference;

KW phage resistance; ds; gene; dased.

OS Synthetic.

XX US2004261149-A1.

XX 23-DEC-2004.

```

XX 24-FEB-2004; 2004US-00785862.
PF 24-FEB-2003; 2003US-0449646P.
PR 24-FEB-2003; 2003US-0449646P.
XX (FAUC/) FAUQUET C M.
PA (PADM/) PADMANABHAN C.
PA (RAMA/) RAMACHANDRAN V.
XX Fauquet CM, Padmanabhan C, Ramachandran V;
XX WPI; 2005-038811/04.
DR 2005-038811/04.
PT Suppressing germiniviral genes in a plant cell comprises introducing a
PT short interfering RNA into the plant cell.
XX Example 1; SEQ ID NO 10; 15pp; English.
XX The invention relates to a novel method for suppressing the expression of
CC a target gene in a plant cell, comprising introducing a short interfering
CC RNA (siRNA) into the plant cell. The method for suppressing the
CC expression of a target gene in a plant cell comprises: introducing a
CC siRNA into the plant cell, where the siRNA is a double-stranded molecule
CC with a first strand consisting of a nucleic acid sequence which is
CC substantially complementary to a nucleic acid sequence of a target gene
CC and a second strand consisting of a nucleic acid sequence which is
CC substantially complementary to the first strand, where the siRNA is
CC capable of suppressing the expression of the target gene. The method,
CC siRNA and composition are useful for suppressing expression of a target
CC gene in a plant cell. Particularly, they are useful for suppressing or
CC inhibiting viral gene expression, specifically germiniviral gene
CC expression, in plant cells. In a particular embodiment of the invention,
CC siRNAs are targeted to the coding region (ACI) of the replication-
CC associated protein (Rep) of an ACMV-CM strain. This polynucleotide
CC sequence represents the coding region of a gene used in the plant gene
CC suppression method of the invention.
XX
SQ Sequence 747 BP; 159 A; 248 C; 224 G; 116 T; 0 U; 0 Other;

```

## Alignment Scores:

Pred. No.: 7.78e-137 Length: 747  
 Score: 1186.00 Matches: 219  
 Percent Similarity: 98.22% Conservative: 2  
 Best Local Similarity: 97.33% Mismatches: 4  
 Query Match: 97.69% Indels: 0  
 DB: 14 Gaps: 0

US-10-006-922A-12 (1-225) x ADV96699 (1-747)

```

QY 1 MetArgSerSerLybAnVal11leLybGluPhenMetArgPheLybValArgMetGluGly 20
DB 1 ATGGCTCTCTCCGAAAGCATCATCACGAGTTCAATGCGCTTCAAGGTCGATGAGGCG 60
QY 21 ThrValAsnGlyH1sgLupHegLui1egLugLyGluGlyArgProTyrgLugLy 40
DB 61 ACCGTGAACGGCCACGAGTTGAGATCGAGGCCAGGGCCGAGGCCCGCCCTTACAGAGGC 120
QY 41 HisAnThrValLybLeuLybVal1ThrLybGlyGlyProLeuProPheAlaTrpAsp1le 60
DB 121 CACAAACACCGTGAAGCTGAAAGTGAACCAAGGGGGGCGCCCTGCTTCGCTGGAGACATC 180
QY 61 LeuSerProGlnPheGlnTyrgLySerLybValTyrgValLybHisProAlaAsp1lePro 80
DB 181 CTGTCCCCCGCCAGTTCAGATGAGCTCCAAAGGTGAAGTGAACCCCGCGACATCCCC 240
QY 81 AspTyrgLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100
DB 241 GACTACAAAGAACTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTATGAATCTTGAG 300
QY 101 AspGlyGlyValVal1ThrVal1ThrGlnAspSerSerLeuGlnAspGlyCyaspHe1leTy 120
DB 301 GACGCGCGCGTGGCGACCGTGAACCAAGGACTCTCCCTCGCAGAGACGGCTGCTCATCTAC 360

```

Qy 121 LyValValyVpheiIleGlyValAlaSpnPhpSerAapGlyProValMetGlnLysLeuThr 140  
 Db 361 AAGGTGAAGTTTCATCGCGGTGAACCTCCCTCCGACGGCCCCGGATGACAGAAAGACC 420  
 Qy 141 MetGlyTTPGluAlaSerThGluArgLeuTyProAaAapGlyValLysLeuGlyGlu 160  
 Db 421 ATGGGCTGGAGGCTCCACGAGCGCTGTACCCCGCGAGCGGTCTAAGGCGGAG 480  
 Qy 161 IleHisValAlaLeuLysLeuLysAapGlyGlyHisTyLeuValGluPhelySerIle 180  
 Db 481 ACCCAAGAGCCCTGAAGCTTAAGACGCGGCCACTCTCTGTGGAGTTCAAGTCCATC 540  
 Qy 181 TyMetAlaLysLysPProValGlnLeuProGlyTyTyTyTyTyValAapSerLysLeuAap 200  
 Db 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCTACTACTACGTGGAGCAAGCCAGGCGGAC 600  
 Qy 201 IleThrSerHisAaenGluAapTyThrIleValGluGlnTyTyTyTyTyTyGluArgThGluGlyArg 220  
 Db 601 ATCACTCTCCCAACAGAGACTACACATCGTGGAGCACTACAGAGCAAGGAGGCGCC 660  
 Qy 221 HisHisLeuPheLeu 225  
 Db 661 CACCACTCTTCTCTG 675

RESULT 55  
 ADM43589  
 ID ADM43589 standard; DNA; 4200 BP.  
 AC ADM43589;  
 XX 07-APR-2005 (first entry)  
 XX DB UBCP-lox-miRMSOD2-BGFP-lox-RFP.  
 XX 89; hairpin; Neuroprotective; shRNA; miRNA; aberrant gene function;  
 KM gain-of-function mutation; neurological disease.  
 XX OS Synthetic.  
 XX WO2005007877-A2.  
 XX 27-JAN-2005.  
 XX 19-JUL-2004; 2004WO-US023789.  
 XX 18-JUL-2003; 2003US-0488510P.  
 XX (UYMA-) UNIV MASSACHUSETTS.  
 XX Xu Z, Xia X;  
 PI WPI; 2005-112883/12.  
 XX Example 6; SEQ ID NO 9; 88bp; English.

This sequence represents the construct UBCP-lox-miRMSOD2-BGFP-lox-RFP. In this construct, the Ubiquitin C promoter (UBCP) normally directs the synthesis of the hairpin and EGFP. However, upon exposure to Cre, the hairpin and EGFP are excised thus preventing hairpin expression, while the RFP gene is under the control of the UBCP. This sequence is a complementary vector to UBCP-lox-RFP-lox-miRMSOD2-BGFP. The construct of the invention comprises a U6 promoter operably linked to a shRNA encoding nucleic acid sequence. The construct further comprises a first loxP site upstream of the promoter and a second loxP site downstream of the shRNA encoding sequence, the loxP sites are in the same orientation so that the promoter and encoding sequences are excisable upon exposure to Cre. The shRNA encoding sequence comprises a first stem-encoding portion, a loop-encoding portion, and a second stem-encoding portion, the loop-encoding portion comprising a first loxP site operably linked to a transcription

CC termination signal upstream of a spacer DNA and a second loxP site, the  
 CC loxP sites are in the same orientation so that the first loxP site, the  
 CC termination signal and spacer DNA are excisable upon exposure to Cre. The  
 CC construct further comprises spacer DNA downstream of the shRNA encoding  
 CC sequence, a second loxP site downstream of the spacer DNA, and a first  
 CC loxP site within the loop-encoding portion of the shRNA encoding  
 CC sequence, the loxP sites are in the same orientation so that the spacer  
 CC DNA and second stem-encoding sequence are excisable upon exposure to Cre.  
 CC The U6 promoter comprises a distal sequence element (DSE), a proximal  
 CC sequence element (PSE), and a TATA box, operably linked, where the  
 CC construct comprises a first loxP site downstream of the shRNA encoding  
 CC sequence, and a second loxP site between the DSE and the PSE, the loxP  
 CC sites are in the same orientation so that the shRNA encoding sequences  
 CC and a portion of the promoter comprising the PSE and the TATA box are  
 CC excisable upon exposure to Cre. The construct also comprises a first loxP  
 CC site and a second loxP site, the sites are interrupted by spacer DNA,  
 CC between the DSE and the PSE, the loxP sites are in the same orientation  
 CC so that a loxP site and the spacer DNA are excisable upon exposure to  
 CC Cre. The construct is useful for the treatment of a disease, which is  
 CC caused by aberrant gene function. The disease is a dominant, gain-of-  
 CC function mutation, e.g. a neurological disease. The constructs are useful  
 CC for regulating shRNA expression in cells or animals. The constructs,  
 CC methods, kits, and compositions are useful for investigating gene  
 CC functions, both physiologic and pathologic, in specific cell groups and  
 CC in specific ages, in normal and diseased pathways.

XX  
 XX Sequence 4200 BP; 889 A; 1094 C; 1265 G; 950 T; 0 U; 2 Other;  
 SQ

Alignment Scores:  
 Pred. No.: 8,85e-136 Length: 4200  
 Score: 1186.00 Matches: 219  
 Percent Similarity: 98.22% Conservative: 2  
 Best Local Similarity: 97.33% Mismatches: 4  
 Query Match: 97.69% Indels: 0  
 DB: 14 Gaps: 0

US-10-006-922a-12 (1-225) x ADM43589 (1-4200)

Qy 1 MetArgSerSerLysAaenValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
 Db 3267 ATGGCTCTCTCCGAGAACGTCAATCAACGATTCATGCCCTTCAAGTGCATGAGAGGAC 3326  
 Qy 21 ThrValAaenGlyHisGluPheGlnIleGluGlyGluGlyGlyArgProTyGluGly 40  
 Db 3327 ACCGTGAACGGCCACGAATTCGAGATCGAGGCGAGGCGAGGCGCCCTTACAGAGGC 3386  
 Qy 41 HisAaenThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAapIle 60  
 Db 3387 CACAAACCGCTGAAGCTTAAGTGAACAAAGGCGGCCCTTGCCTTGGCTGGGACATC 3446  
 Qy 61 LeuSerProGlnPheGlnTyGlySerLysValTyValLysHisProAlaAapIlePro 80  
 Db 3447 CTGTCCCGCCAGTTCCAGTACGAGCTCCCAAGGTGTACGTGAAGCAACCCCGCAGACATCCC 3506  
 Qy 81 AapTyTyLysLysLeuSerPheProGluGlyPheLysTyTTPGluArgValMetAaenPheGlu 100  
 Db 3507 GACTACAAAGAGCTGTCTCTCCCGAGGACCTTCAAGTGAAGCGGTGAATTAATCTTCAG 3566  
 Qy 101 AapGlyGlyValAlaThrValThrGlnAapSerSerLeuGlnAapGlyCysPheIleTy 120  
 Db 3567 GACGGCGGCGTGGCAACGTACCAAGACTCTCTCCGAGGAGCGGCTGTCATCTAC 3626  
 Qy 121 LyValValyVpheiIleGlyValAlaSpnPhpSerAapGlyProValMetGlnLysLeuThr 140  
 Db 3627 AAGGTGAAGTTTCATCGCGGTGAACCTCCCTCCGACGGCCCCGGATGACAGAAAGACC 3686  
 Qy 141 MetGlyTTPGluAlaSerThGluArgLeuTyProAaAapGlyValLysLeuGlyGlu 160  
 Db 3687 ATGGGCTGGAGGCTCCACGAGCGCTGTACCCCGCGAGCGGTCTAAGGCGGAG 3746  
 Qy 161 IleHisValAlaLeuLysLeuLysAapGlyGlyHisTyLeuValGluPhelySerIle 180  
 Db 3747 ACCCAAGAGCCCTGAAGCTTAAGACGCGGCCACTCTCTGTGGAGCAAGTCCATC 3806

QY 181 TyMetAlaIySblyPProValGlnLeuProGlyTyrTyrTyrValaAspSerLyLeuAsp 200  
DB 3807 TACATGGCCAAAGAGCCCGTGCACCTGCCCGGCTACTACTACGTGAGCGCCAACTGAC 3866  
QY 201 lIethSerHsBaNgIuAapPyrThrIleValGlnGlnTyrGluAqTThGluGlyArg 220  
DB 3867 ATCACTCTCCCAACGAGGACTACACATCTGTGAGCATGACGAGCCGACCGAGGGCCGC 3926  
QY 221 HIsHIsleupheLau 225  
DB 3927 CACCACCTGTTCTCG 3941  
RESULT 56  
ADM43588  
ID ADM43588 standard; DNA; 4300 BP.  
XX  
AC ADM43588;  
XX  
DT 07-APR-2005 (first entry)  
XX  
DE UDCP-lox-RFP-lox-miRMSOD2-EGFP.  
XX  
KM sb; hairpin; Neuroprotective; shRNA; miRNA; aberrant gene function;  
KW gain-of-function mutation; neurological disease.  
XX  
OS Synthetic.  
XX  
SN WO200507877-A2.  
XX  
PN 27-JAN-2005.  
PD  
XX 19-JUL-2004; 2004WO-US023789.  
XX 18-JUL-2003; 2003US-0488510P.  
PR  
XX (UWMA-) UNIV MASSACHUSETTS.  
PA  
XX Xu Z, Xia X;  
XX  
XX WPI; 2005-112883/12.  
DR  
XX  
PT New construct comprises a U6 promoter operably linked to a small hairpin  
PT RNA (shRNA) encoding nucleic acid sequence, useful for regulating shRNA  
PT expression in cells or animals.  
XX  
XX  
XX Example 6; SEQ ID NO 8; 88bp; English.  
XX  
CC This sequence represents the construct UDCP-lox-RFP-lox-miRMSOD2-EGFP. In  
CC this construct, the Ubiquitin C promoter (UBCP) normally transcribes RFP  
CC mRNA, which terminates at a polyA site. However, because the RFP gene is  
CC flanked by two loxP sites, upon exposure to cre, the RFP gene is excised  
CC and the construct is converted to UDCP-lox-miRMSOD2-EGFP. The resulting  
CC recombinant sequence transcribes a message that contains an intron with an  
CC shRNA (small hairpin) against the mouse SOD2 gene and an EGFP gene. The  
CC construct of the invention comprises a U6 promoter operably linked to a  
CC shRNA encoding nucleic acid sequence. The construct further comprises a  
CC first loxP site upstream of the promoter and a second loxP site  
CC downstream of the shRNA encoding sequence, the loxP sites are in the same  
CC orientation so that the promoter and encoding sequences are excisable  
CC upon exposure to Cre. The shRNA encoding sequence comprises a first stem-  
CC encoding portion, a loop-encoding portion, and a second stem-encoding  
CC portion, the loop-encoding portion comprising a first loxP site operably  
CC linked to a transcription termination signal upstream of a spacer DNA and  
CC a second loxP site, the loxP sites are in the same orientation so that  
CC the first loxP site, termination signal and spacer DNA are excisable upon  
CC exposure to Cre. The construct further comprises spacer DNA downstream of  
CC the shRNA encoding sequence, a second loxP site downstream of the spacer  
CC DNA, and a first loxP site within the loop-encoding portion of the shRNA  
CC encoding sequence, the loxP sites are in the same orientation so that the  
CC spacer DNA and second stem-encoding sequence are excisable upon exposure  
CC to Cre. The U6 promoter comprises a distal sequence element (DSB), a  
CC proximal sequence element (PSE), and a TATA box, operably linked, where

CC the construct comprises a first loxP site downstream of the shRNA  
CC encoding sequence, and a second loxP site between the DSB and the PSE.  
CC The loxP sites are in the same orientation so that the shRNA encoding  
CC sequences and a portion of the promoter comprising the PSE and the TATA  
CC box are excisable upon exposure to Cre. The construct also comprises a  
CC first loxP site and a second loxP site, the sites are interrupted by  
CC spacer DNA, between the DSB and the PSE, the loxP sites are in the same  
CC orientation so that a loxP site and the spacer DNA are excisable upon  
CC exposure to Cre. The construct is useful for the treatment of a disease,  
CC which is caused by aberrant gene function. The disease is a dominant,  
CC gain-of-function mutation, e.g. a neurological disease. The constructs  
CC are useful for regulating shRNA expression in cells or animals. The  
CC constructs, methods, kits, and compositions are useful for investigating  
CC gene functions, both physiologic and pathologic, in specific cell groups  
CC and in specific ages, in normal and diseased pathways.  
XX  
SQ Sequence 4300 BP; 906 A; 1128 C; 1297 G; 967 T; 0 U; 2 Other;  
Alignment Scores:  
Pred. No.: 9.15e-136 Length: 4300  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 14 Gaps: 0  
US-10-006-922A-12 (1-225) x ADM43588 (1-4300)  
QY 1 MetArgSerSerLySbaNValIleLyGluPheMetArgPheLySValaArgMetGluGly 20  
DB 1291 ATGGCCCTCTCCGAGAAAGTCATCACCGAGTTCATGCTTCAAGGCGCCATGAGAGGC 1350  
QY 21 ThrValaAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 1351 ACCGTGAACGCGCAAGGTTCTGAGATCGAGGGCGAGGCGCCGCTTACAGAGGC 1410  
QY 41 HisAsnThrValIlySLeuLySValaThrLyGlyGlyProLeuProPhealaTyrAspIle 60  
DB 1411 CACACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCTGCTTCCGTGGAGCATC 1470  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValaTyrValIlySHisProAlaAspIlePro 80  
DB 1471 CTGTCCTCCCGAGTTCCTCGATCGAGCTCCAGGTGACGTGAAGCACCCCGCGACATCCCC 1530  
QY 81 AspTyrTyrLySLeuSerPheProGluGlyPheLySProGluValMetAsnPheGlu 100  
DB 1531 GACTACAAAGAGCTGTCTTCCCGAGGCGCTTCAAGTGGAGCGCGTGAATGACTTGAG 1590  
QY 101 AspGlyGlyValaValaThrValaThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120  
DB 1591 GACGCGCGCGTGGAGACCGTGACCCAGGACTCTCTCCGTGAGAGCGGCTGTTCACTGAC 1650  
QY 121 LyValIlySbHeIleGlyValaAsnPheProSerAspGlyProValMetGlnLySbThr 140  
DB 1651 AAGGTGAAGTTCACTCGCGTGAACTTCCCTCCGACGCGCCCGTGAAGCAGAAAGAACCC 1710  
QY 141 MetGlyTyrProGluIlySbSerThrGluValArgLeuTyrProArgAspGlyValaIlySb 160  
DB 1711 ATGGGCTGGAGAGGCTTCCACCGAGCGCTTACCCCGCAGCGCGTGAAGAGGAGAG 1770  
QY 161 lIethIlySbAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
DB 1771 ACCCAAGAGCCCTGAAGCTGAAGAGAGGCGGCACTACTGTGTGAGTTCAATCTCATC 1830  
QY 181 TyMetAlaIySblyPProValGlnLeuProGlyTyrTyrTyrValaAspSerLyLeuAsp 200  
DB 1831 TACATGGCCAAAGAGCCCGTGCACCTGCCCGGCTACTACTACGTGAGCGCCAACTGAC 1890  
QY 201 lIethSerHsBaNgIuAapPyrThrIleValGlnGlnTyrGluAqTThGluGlyArg 220  
DB 1891 ATCACTCTCCCAACGAGGACTACACATCTGTGAGCATGACGAGCCGACCGAGGGCCGC 1950  
QY 221 HIsHIsleupheLau 225

Db 1951 CACCACCTGTTCTTG 1965

RESULT 57

ADQ91220 standard; DNA; 4488 BP.

ADQ91220/  
04-NOV-2004 (first entry)

Fusion construct of human DCX promoter and the DsRed2 gene SeqID 6.

human; neuronal marker; doublecortin; DCX; ds; DsRed2; chimeric; bcl-2; PTEN; dopaminergic determination factor; Nurr1; neuronal migration; neurotrophin; SDF-1; NeuroD; BMP4; transgenic; neurogenic; neural stem cell activity; neuronal differentiation; neurodegenerative disorder; epilepsy; stroke; learning; memory skill; neurodegenerative disorder; Parkinson's; Huntington's; HIV-dementia; psychiatric disorder; tumorous disorder; brain malformation; vasotropic; anticonvulsant; cerebroprotective; antiparkinsonian; neuroprotective; neurotropic; anti-HIV; regulatory sequence; proliferative neuronal determined cell; anti-apoptotic.

Homo sapiens.  
Chimeric.  
Unidentified.

MO2004067751-A1.

12-AUG-2004.

28-JAN-2004; 2004MO-EP000760.

28-JAN-2003; 2003BP-00002027.

(UYRE-) UNIV KLINIKUM REGENSBURG.

Couillard-Despres S, Karl C, Kuhn H, Aigner L;

WPI: 2004-581007/56.

Use of a regulatory sequence for the early, transient expression of a heterologous nucleotide sequence in proliferative neuronal determined cells.

Claim 20; SEQ ID NO 6; 270bp; English.

This invention relates to the novel use of regulatory sequences for the early, transient expression of a heterologous nucleotide sequences in proliferative neuronal determined cells. Specifically, it refers to the preparation of a recombinant vector comprising a heterologous gene sequence such as an anti-apoptotic gene e.g. bcl-2 or PTEN, a determination/differentiation gene such as the dopaminergic determination factor Nurr1, a gene capable of inducing neuronal migration e.g. neurotrophin or SDF-1 or a transcription factor gene such as NeuroD or BMP4. The present invention describes non-human transgenic animals that have been transfected with such a vector that can be used for screening compounds in vitro to identify those with neurogenic properties capable of regulating neural stem cell activity, neurogenesis and/or neuronal differentiation, which can subsequently be used in the development of appropriate pharmaceutical compositions. Furthermore, these transgenic animals can be used for the in vivo tracking of newly generated neurons, as well as for transplanted or migrating neuronal determined cells. Accordingly, such compositions can be used in gene therapy for neurological disorders e.g. epilepsy or stroke and for the enhancement of learning and/or memory skills caused by neurodegenerative disorders such as Parkinson's, Huntington's or HIV-dementia, as well as psychiatric disorders, tumorous disorders of the nervous system or a developmental or brain malformation. Hence, these compositions exhibit vasotropic, anticonvulsant, cerebroprotective, antiparkinsonian, neuroprotective, neurotropic and anti-HIV activities. This polynucleotide sequence is the fusion construct DNA of the reporter gene DsRed2 under the control of the

CC human neuronal marker doublecortin (DCX) promoter, given in an CC exemplification of the invention.

XX Sequence 4488 BP; 1184 A; 1026 C; 973 G; 1305 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9,72e-136	Length:	4488
Score:	1186.00	Matches:	219
Percent Similarity:	98.22%	Conservative:	2
Best Local Similarity:	97.33%	Mismatches:	4
Query Match:	97.69%	Indels:	0
DB:	13	Gaps:	0

US-10-006-922a-12 (1-225) x ADQ91220 (1-4488)

QY	1	MetArgSerSerLyVaenValIleLygluPheMetArgPheLyValArgMetGluGly	20
DB	3572	ATGGCCTTCCTCCGAAAGTCATCACCAGATTCAATGCGCTTCAAGTGCATGAGAGGC	3631
QY	21	ThrValAsnGlyVHAsgluPheGluIleGluGlyVgGluGlyArgProTyrgGluGly	40
DB	3632	ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGCGCTTACAGAGGC	3691
QY	41	HIsAnThrValLyLeuLyVaIThrLyAsGlyGlyProLeuProPheAlaTrpAspIle	60
DB	3692	CACAAACCGTGAAGCTGAAGGTACCAAGGCGGCCCTTCCCTTGCCTGGAGATC	3751
QY	61	LeuSerProGlnPheGlnTyrgLySerLyVaIlyrValLySHsAProAlaAspIlePro	80
DB	3752	CTGTCCCGCCAGTTCAGATACGGCTCCAAAGGTATACGAAACACCCCGCAATCCCC	3811
QY	81	AspTyrlLyLyLeuSerPheProGluGlyPheLySTrPGluVgValMetAsnPheGlu	100
DB	3812	GACTTAAAGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATAACTTCAG	3871
QY	101	AspGlyGlyValIleThrValIThrGlnAspSerSerLeuGlnAspGlyCySPheIleTy	120
DB	3872	GACGGCGCGTGGCGACCTGATCCAGACTCTCTCCGACGAGACGCTGCTTCACTAC	3931
QY	121	LySValLySPheIleGlyValIAsnPheProSerAspGlyProValMetGlnLyAlThr	140
DB	3932	AAAGTGAAAGTTCATCGGCGGTGAATCTCCCTCCGACGGCCCGTGAACAGAAAGACC	3991
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyProAlaAspGlyValIleuLyAsGlu	160
DB	3992	ATGGGCTGGAGGCTTCCACGAGCGCTGTACCCCGCAGCGGTGCTGAAGGCGCAG	4051
QY	161	IleHsILySValIleuLyLeuLyAspGlyGlyHsTyrlLyLeuValIgluPheLySerIle	180
DB	4052	ACCCCAAGGCGCTGAAGCTGAAGACGCGGCCACTACCTGTGAGATTCAAGTCAATC	4111
QY	181	TyrMetAlaLyLySProValGlnLeuProGlyTyrgTyrgValAspSerLyLeuAsp	200
DB	4112	TACATGGCCAAAGAACCGGTGAGTCCCGGCTACTACTAGTGAAGCCCAAGCTGAGC	4171
QY	201	IleThrSerHsILyAsnGluAspTyrlThrIleValGluGlnTyrgGluArgThrGluGly	220
DB	4172	ATCACCTCCCAAGAGAGACTTACATCGTGAAGAGTACGAGCGCACCGAGGCGCGC	4231
QY	221	HIsHsILyLeuPheLeu 225	
DB	4232	CACCACCTGTTCTTG 4246	

RESULT 58

ADL35208 standard; DNA; 6706 BP.

ADL35208;

03-JUN-2004 (first entry)

Plasmid pL3.2 nucleotide sequence SEQ ID NO:4.



QY 181 TyMetAlaLyLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLyLeuAsp 200  
 DB 3318 TACATGGCCAAAGAGCCGCTGACGCTGCTACTACTACGTGAGACGCCAAGCTGGAC 3377  
 QY 201 ILeThSerHiAaenGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
 DB 3378 ATCACTCCCAACAGAGACTACACATCGTGAGAGTACAGGCGACCGAGGCGCGC 3437  
 QY 221 HisHLeuPheLeu 225  
 DB 3438 CACCACTGTTCTCTG 3452  
 RESULT 59  
 ADR30815  
 ID ADR30815 standard; DNA; 7495 BP.  
 AC ADR30815;  
 XX 04-NOV-2004 (first entry)  
 DT Zebrafish BDNF gene construct DNA #3.  
 DE Zebrafish BDNF gene construct DNA #3.  
 XX Brain-derived neurotrophic factor; BDNF; promoter; zebrafish, de.  
 KM Danio rerio.  
 OS Chimeric.  
 OS Unidentified.  
 FT Key Location/Qualifiers  
 FT misc\_feature 1..20  
 FT /\*tag= a  
 FT /note= "Vector DNA"  
 FT misc\_feature 15..5091  
 FT /\*tag= b  
 FT /note= "Expression vector DNA"  
 FT misc\_feature 21..1776  
 FT /\*tag= c  
 FT /note= "Zebrafish BDNF gene 5' flank region"  
 FT promoter 1777..1804  
 FT /\*tag= d  
 FT /note= "Zebrafish BDNF promoter gene"  
 FT exon 1805..2099  
 FT /\*tag= e  
 FT /note= "Exon 1c (5' UT) of Zebrafish BDNF gene"  
 FT misc\_feature 2100..2122  
 FT /\*tag= f  
 FT /note= "Vector DNA"  
 FT misc\_feature 2123..2800  
 FT /\*tag= g  
 FT /note= "Reporter vector DNA"  
 FT misc\_feature 2801..2807  
 FT /\*tag= h  
 FT /note= "Linker DNA"  
 FT misc\_feature 2808..5086  
 FT /\*tag= i  
 FT /note= "3' flank region"  
 FT misc\_feature 5087..7495  
 FT /\*tag= j  
 FT /note= "Vector DNA"  
 FT US2004157294-A1.  
 PD 12-AUG-2004.  
 XX 23-DEC-2003; 2003US-00742828.  
 XX 30-NOV-2001; 2001US-0334079P.  
 PR 29-NOV-2002; 2002US-00306737.  
 XX (USGO ) US DEPT VETERANS AFFAIRS.  
 PA Heinrich G, Huynh G;  
 PI XX

DR WP1: 2004-592772/57.  
 XX Nucleic acid construct useful for screening substances for their ability  
 PT to influence production of brain derived neurotrophic factor (BDNF),  
 PT comprises promoter of BDNF gene and red fluorescent reporter.  
 PS Claim 26; SEQ ID NO 5; 56pp; English.  
 XX The present invention provides an isolated brain-derived neurotrophic  
 CC factor (BDNF) gene promoter and its associated nucleic acid construct.  
 CC The invention is useful for screening a number of substances for their  
 CC ability to influence the production of BDNF in living organisms. The  
 CC present sequence is zebrafish brain-derived neurotrophic factor (BDNF)  
 CC gene construct DNA. This sequence is used in the invention.  
 XX Sequence 7495 BP; 2249 A; 1599 C; 1659 G; 1988 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2e-135 Length: 7495  
 Score: 1186.00 Matches: 219  
 Percent Similarity: 98.22% Conservative: 2  
 Best Local Similarity: 97.33% Mismatches: 4  
 Query Match: 97.69% Indels: 0  
 DB: Gaps: 0  
 US-10-006-922a-12 (1-225) x ADR30815 (1-7495)  
 QY 1 MecArgSerLeuValAsnValIleValGluPheMetArgPheValArgMetGluGly 20  
 DB 2123 ATGGCTCTCCGAGAACTGATCATCGATTACGCTTCAAGTGCATGAGAGGC 2182  
 QY 21 ThrValaAngIyHLeGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
 DB 2183 ACCGTGAACGGCCACGAGTTCAGATCGAGGCGGAGGCGGCGCCCTACGAGGAC 2242  
 QY 41 HisAenThrValLeuLeuValThrArgGlyGlyProLeuProPheAlaTyrAspIle 60  
 DB 2243 CACCAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCTGCGGACATC 2302  
 QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80  
 DB 2303 CTGTCCCGCCAGTTCAGATCGAGGCTCAAGGGTGACGAGACCCCGCGACATCCCC 2362  
 QY 81 AspTyrLyLeuSerPheProGluGlyPheLyTyrGluArgValMetAsnPheGlu 100  
 DB 2363 GACTACAAAGAGCTGTCTCCCGAGGAGCTTCAAGTGGAGCGGTGATGAACCTCGAG 2422  
 QY 101 AspGlyGlyValValThrValThrGlnAspSerLeuGlnAspGlyCysPheIleTyr 120  
 DB 2423 GACGGCGGCGTGGCCAGCGTGAACCGACATCTCTCCGACGAGCGGCTTCACTTAC 2482  
 QY 121 LySValLyPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
 DB 2483 AAGGTGAAGTTCATCGGCTGTAACCTTCCCTCCGAGCGGCGGTGATGAGAAAGAAC 2542  
 QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySLeu 160  
 DB 2543 ATGGGCTGGAGGCTCCACCGAGCGGCTGTACCCCGGAGCGGCGTGTGAAGGCGAG 2602  
 QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGlnPheLySLeu 180  
 DB 2603 ACCCAAGAGCCCTTAAGCTGAAGACGGCGGCGACACTACGTGAGATTCAAGCTTATC 2662  
 QY 181 TyMetAlaLyLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLyLeuAsp 200  
 DB 2663 TACATGGCCAAAGAGCCGCTGACGCTGCTACTACTACGTGAGACGCCAAGCTGGAC 2722  
 QY 201 ILeThSerHiAaenGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
 DB 2723 ATCACTCCCAACAGAGACTACACATCGTGAGAGTACAGGCGACCGAGGCGCGC 2782  
 QY 221 HisHLeuPheLeu 225  
 |||||||||||||||



Db 2783 CACCACCTGTTCTCG 2797

RESULT 60

ADQ91222

ID ADQ91222 standard; DNA; 7616 BP.

XX

XX ADQ91222;

XX

DT 04-NOV-2004 (first entry)

DE

XX Expression vector phUDCXPromODSRed2 DNA SeqID 8.

XX human; neuronal marker; doublecortin; DCX; ds; DShed2; expression vector;

KW bcl-2; PTEN; dopaminergic determination factor; Nurr1;

KW neuronal migration; neurophilin; SDF-1; Neurod; BMP4; transgenic;

KW neurogenic; neural stem cell activity; neuronal differentiation;

KW neurodegenerative disorder; epilepsy; stroke; learning; memory skill;

KW neurodegenerative disorder; Parkinson's; Huntington's; HIV-dementia;

KW psychiatric disorder; tumorous disorder; brain malformation; vasotropic;

KW anticonvulsant; cerebroprotective; antiparkinsonian; neuroprotective;

KW neurotropic; anti-HIV; regulatory sequence; anti-apoptotic.

KW proliferative neuronal determined cell; anti-apoptotic.

XX Homo sapiens.

OS Synthetic.

OS Unidentified.

XX MO2004067751-A1.

XX

PD 12-AUG-2004.

XX

PF 28-JAN-2004; 2004MO-EP000760.

XX

PR 28-JAN-2003; 2003EP-00002027.

XX

XX (UYRE-) UNIV KLINIKUM REGENSBURG.

PA

XX Couillard-Despres S, Karl C, Kuhn H, Aigner L;

PI WPI; 2004-581007/56.

DR

XX

PT Use of a regulatory sequence for the early, transient expression of a

PT heterologous nucleotide sequence in proliferative neuronal determined

PT cells.

XX

PS Claim 20; SEQ ID NO 8; 270pp; English.

XX

XX This invention relates to the novel use of regulatory sequences for the

CC early, transient expression of a heterologous nucleotide sequences in

CC proliferative neuronal determined cells. Specifically, it refers to the

CC preparation of a recombinant vector comprising a heterologous gene

CC sequence such as an anti-apoptotic gene e.g. bcl-2 or PTEN, a

CC determination/differentiation gene such as the dopaminergic determination

CC factor Nurr1, a gene capable of inducing neuronal migration e.g.

CC neurophilin or SDF-1 or a transcription factor gene such as Neurod or

CC BMP4. The present invention describes non-human transgenic animals that

CC have been transfected with such a vector that can be used for screening

CC compounds in vitro to identify those with neurogenic properties capable

CC of regulating neural stem cell activity, neurogenesis and/or neuronal

CC differentiation, which can subsequently be used in the development of

CC appropriate pharmaceutical compositions. Furthermore, these transgenic

CC animals can be used for the in vivo tracking of newly generated neurons,

CC as well as for transplanted or migrating neuronal determined cells.

CC Accordingly, such compositions can be used in gene therapy for

CC neurological disorders e.g. epilepsy or stroke and for the enhancement of

CC learning and/or memory skills caused by neurodegenerative disorders such

CC as Parkinson's, Huntington's or HIV-dementia, as well as psychiatric

CC disorders, tumorous disorders of the nervous system or a developmental

CC or brain malformation. Hence, these compositions exhibit vasotropic,

CC anticonvulsant, cerebroprotective, antiparkinsonian, neuroprotective,

CC neurotropic and anti-HIV activities. This polynucleotide sequence is the

CC expression vector phUDCXPromODSRed2, which comprises the reporter gene

CC DShed2 under the control of the human neuronal marker doublecortin (DCX)

CC promoter, given in an exemplification of the invention.

XX

XX Sequence 7616 BP; 1916 A; 1854 C; 1822 G; 2024 T; 0 U; 0 Other;

SO

Alignment Scores:

Pred. No.: 2.05e-135 Length: 7616

Score: 1186.00 Matches: 219

Percent Similarity: 98.22% Conservative: 2

Best local Similarity: 97.33% Mismatches: 4

Query Match: 97.69% Indels: 0

DB: 13 Gaps: 0

US-10-006-922a-12 (1-225) x ADQ91222 (1-7616)

QY 1 MetArgSerSerIysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

DB ATGGCCCTCCCGAAGACGTCATCACCGAATTCATGCGCTTCAGAGTCCGATGAGGCG 3665

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40

DB ACCGTGACGGCCACGAGTTCGAGATGAGGAGGCGAGGCGCGCCCTACGAGGCGC 3725

QY 41 HisAsnThrValLysLysLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60

DB CACAAACACCGTGAAGCTGAAGTGAACCAAGGCGCGCCCTCGCTTCGCTGGAGATC 3785

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

DB CTGTCCCCCGACGTTCCAGTACGGCTCCAGAGTTCAGTGAAGCACCCCGCGACATCCCC 3845

QY 81 AspTyrLysLysLysSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100

DB GACTACAAAGAAAGGCTGCTTCCCGAGGCGCTTCAAGTGGAGCGCGATGAATTCGAG 3905

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

DB GACGCGCGCTGGCGACCGTGAACCGACCTCTCCCTGCGAGGACGCGCTCTCATCTAC 3965

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

DB AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCGACGCGCCCGTGAATGCAAGAAAGCC 4025

QY 141 MetGlyTyrGluLysSerThrGluArgLeuTyrProArgAspGlyValLysLysGlyGlu 160

DB ATGGGCTGGGAGGCTCCACGAGCGCGCTGATCCCCCGAGCGCGCTGAGAGGCGAG 4085

QY 161 IleHisLysAlaLeuLysLysLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

DB ACCCAAGAGCCCTGAAGCTGAAGGAGCGCGCCACTGAGTGGAGTTCAGTCCATC 4145

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

DB TACATGGCCAAAGAGCCCGTGAAGCTTCCCGCTACTACTAGTGAAGCGCAGCTGAGC 4205

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB ATCACTCCCAACAAGAGGACTACACCATCTGTGAGCAGTACGAGGCCACGAGGCGCGC 4265

QY 4266 CACCACCTGTTCTCG 4280

DB

RESULT 61

ADL35211

ID ADL35211 standard; DNA; 7927 BP.

XX

XX ADL35211;

XX

DT 03-JUN-2004 (first entry)

DE

XX Plasmid pL1.5 nucleotide sequence SEQ ID NO:7.

XX

KW lentiviral vector; functional packaging signal; multiple cloning site;





```
Db 4479 ACCCAAGAGCCCTGGAAGCTGAAGAGCGGCCCACTACCTGTCGAGTTCACATCCATC 4538
QY 181 TyTMeTAlAlYsYsPProVAlGInLeuProGlyTyrTyTyrValAAsPSeRlySLeuAAsP 200
Db 4539 TACATGGCCAAAGAGCCCGTCGACCTGCCCGCTACTACTACTAGTGAGCGCAAGCTGAC 4598
QY 201 lLeThSeRtHsAeNgIuAAsPtyrThrIleValGInGInTyrgIuAaGtThGInGlyAaG 220
Db 4599 ATCACTCTCCCAAGAGAGACTACACATCGTGAGAGTACGAGCGCACCGAGGCGCGC 4658
QY 221 HIsHIsLeuPheLeu 225
Db 4659 CACCACTGTTCCTG 4673

RESULT 62
ADM82791
ID ADM82791 standard; cDNA; 9731 BP.
XX
AC ADM82791;
XX
DT 03-JUN-2004 (first entry)
DE DNA repair pathway related retroviral vector cDNA with CMV promoter.
XX
KM inducer; inhibitor; DNA repair pathway; anti-HIV; cytosratic; vinticide;
KM antidiabetic; neuroprotective; retroviral infection; AIDS; HIV infection;
KM cancer; human adult T-cell leukemia; lymphoma;
KM feline immunodeficiency virus; Type I diabetes; multiple sclerosis;
KM gene therapy; human; cyclic; circular; CMV promoter; ss.
XX
OS unidentified.
XX
FN WO2003089573-A2.
XX
PD 30-OCT-2003.
PF 04-APR-2003; 2003WO-US010302.
PR 05-APR-2002; 2002US-0370376P.
XX
PA (FISH/) FISHEL R A.
PA (YODE/) YODER K E.
XX
PI Fishel RA, Yoder KE;
XX
DR WPI; 2003-854096/79.
XX
PT Screening for compounds that modulate a DNA repair pathway and/or
PT retroviral integration, useful for treating retroviral infection,
PT comprises determining the amount of a retroviral cDNA circularization in
PT the presence of the test compound.
XX
PS Claim 73; SEQ ID NO 5; 89pp; English.
XX
CC The invention relates to a novel method for screening for inducers or
CC inhibitors of a DNA repair pathway by contacting at least one component
CC of a DNA repair pathway with a non-circularized retroviral cDNA in the
CC presence and absence of a test compound, and determining whether
CC circularization of the cDNA is increased or decreased in the presence of
CC the test compound. The DNA repair pathway components have the following
CC activities: anti-HIV, cytosratic, vinticide, antidiabetic, and
CC neuroprotective. The method is useful for identifying compounds that
CC modulate a DNA repair pathway and/or retroviral activity. The compound is
CC used in manufacturing a pharmaceutical composition for the treatment of a
CC retroviral infection (e.g. AIDS, HIV infection, cancer, human adult T-
CC cell leukemia, lymphoma, feline immunodeficiency virus, Type I diabetes
CC or multiple sclerosis) or for increasing the efficiency of gene delivery
CC in a gene therapy. This polynucleotide represents a retroviral cDNA
CC sequence of the invention.
XX
SQ Sequence 9731 BP; 2444 A; 2412 C; 2548 G; 2327 T; 0 U; 0 Other;
XX
Alignment Scores:
```

```
Pred. No.: 2,89e-135 Length: 9731
Score: 1186.00 Matches: 219
Percent Similarity: 98.22% Conservative: 2
Best Local Similarity: 97.33% Mismatches: 4
Query Match: 97.69% Indels: 0
DB: 11 Gaps: 0
US-10-006-922a-12 (1-225) x ADM82791 (1-9731)
QY 1 MetASeRSeRlySbAnValIleYbGluPheMeTArGpHeLybValAaGMeTGlUgLy 20
Db 5496 ATGGCTCTCCGAGAAAGTCATCACCGAGTTCATGCGCTTCAAGGTCGCAAGAGAGC 5555
QY 21 ThrValAaNgLyHsGluPheGluIleGluGlyGluGlyGluGlyAaGProTyrgIuGly 40
Db 5556 ACCGTGAACGGCCACGAGTTCGAGATTCGAGAGGCGAGAGGCGCCCTTACGAGAGGC 5615
QY 41 HIsAnThrValLybLeuLybValThrLyGlyGlyProLeuProPheAlATPAsPile 60
Db 5616 CACACACCGTGAAGCTGAAGGTGACCAAGGCGCGCCCTTCCCTTCCCTGGAGATC 5675
QY 61 LeuSeRProGInPheGInTyrgIySeRlybValTyTyrValIyShIsPProAlAAsPilePro 80
Db 5676 CTGTCCTCCCGAGTTCAGTACGAGCTCCAGAGTGTACGTGAAGCACCCGCGACATCCC 5735
QY 81 AsPTyTlybLySeLeuSeRPhEProGInGlyPhElbETrGluArgValMeTAnPheGlu 100
Db 5736 GACTACAAAGAGCTGCTCTCCCGAGGCTTCAAGTGGAGGCGCGATGAACCTTGAG 5795
QY 101 AsPgLyGlyValValThrValThrGInAAsPSeRSeRleuGInAAsPgLyCybPheIleTyR 120
Db 5796 GACGCGCGCTTGCGAGACCGTGACCCAGACTCTCTCCGTGAGAGACGCTTCAATCATC 5855
QY 121 LybValLybPheIleGlyValAaPheProSeRAsPgLyProValMeTGlNlybLyEThr 140
Db 5856 AAGGTGAAGTTCATCGCGCTGAACCTTCCCTCCGACGCGCCCGTGAATGACAGAAAGAC 5915
QY 141 MeTGlTyTProGluASeRThrGluArgLeuTyProArgAsPgLyValLeuLybGlyGlu 160
Db 5916 ATGGCTCGGAGGCTTCACCGAGCGCTGTACCCCGCAGCGCGCTGAGAGGGGAG 5975
QY 161 lLeHIsbYsAlaLeuLybLeuLybAAsPgLyGlyHIsTyRLeuValGluPheLySeRile 180
Db 5976 ACCCAAGAGCCCTGAAGCTGAAGAGCGGCGCACACTCTGTGTGAGTTCAGTCCATC 6035
QY 181 TyTMeTAlAlYsYsPProVAlGInLeuProGlyTyrTyTyrValAAsPSeRlySLeuAAsP 200
Db 6036 TACATGGCCAAAGAGCCCGTCGACCTGCCCGCTACTACTACTAGTGAGCGCAAGCTGAC 6095
QY 201 lLeThSeRtHsAeNgIuAAsPtyrThrIleValGInGInTyrgIuAaGtThGInGlyAaG 220
Db 6096 ATCACTCTCCCAAGAGAGACTACACCATCTGTGAGAGTACGAGCGCACCGAGGCGCGC 6155
QY 221 HIsHIsLeuPheLeu 225
Db 6156 CACCACTGTTCCTG 6170

RESULT 63
ADM82792
ID ADM82792 standard; cDNA; 9782 BP.
XX
AC ADM82792;
XX
DT 03-JUN-2004 (first entry)
DE DNA repair pathway related retroviral vector cDNA with MSH2 promoter.
XX
KM inducer; inhibitor; DNA repair pathway; anti-HIV; cytosratic; vinticide;
KM antidiabetic; neuroprotective; retroviral infection; AIDS; HIV infection;
KM cancer; human adult T-cell leukemia; lymphoma;
KM feline immunodeficiency virus; Type I diabetes; multiple sclerosis;
KM gene therapy; human; cyclic; circular; MSH2 promoter; ss.
XX
```



```

XX The present sequence is the cDNA sequence of drFP583 from Dicosoma sp.
CC "red", a non-bioluminescent species of the Class Anthozoa. drFP583 is a
CC full-length cDNA encoding a novel fluorescent protein (NFP). Fluorescent
CC proteins can be used in fluorescent labeling, a useful tool for marking a
CC protein, cell or organism of interest. Unlike other markers used in
CC protein labeling, such as beta-galactosidase and luciferase, fluorescent
CC proteins do not require an exogenous cofactor or substrate. Methods
CC involving fluorescent proteins are also less laborious and less difficult
CC to control than the traditional methods of fluorescent labeling, where a
CC protein of interest is purified and then covalently conjugated to a
CC fluorophore derivative. Novel fluorescent proteins isolated from species
CC of the Class Anthozoa can be used as markers for gene expression and
CC protein localization studies, and in fluorescence resonance energy
CC transfer (FRET) reactions. They may have improved properties and better
CC suitability for larger excitations compared to prior art fluorescent
CC proteins such as green fluorescent protein. (Updated on 15-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 898 BP; 273 A; 172 C; 240 G; 213 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.78e-136 Length: 898
Score: 1184.00 Matches: 219
Percent Similarity: 98.22% Conservative: 2
Best Local Similarity: 97.33% Mismatches: 4
Query Match: 97.53% Indels: 0
Gaps: 0
US-10-006-922A-12 (1-225) x AAA48742 (1-898)
QY 1 MetPArgSerSerIysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db 93 ATGAGGCTTCCTCAAGAAAGTTATCAAGAGATTCAAGGTTTAAGTTCGATGAAGGA 152
QY 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 40
Db 153 ACGGTCATATGGGACGACGAGTTGAAATAGAAAGCCAAAGAGAGGAGCCATACGAAGGC 212
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 213 CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTTGGCTTGGGATATT 272
QY 61 LeuSerProGlnPheGlnIleGlySerLysValLysValLysValLysValLysVal 80
Db 273 TTGTCCACCAATTCAGATGGAAGCAAGGTAATATGCAAGCACCCGCGCAATACCA 332
QY 81 AspTyrLysLysLeuSerPheProGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 100
Db 333 GACATATAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACTTTGA 392
QY 101 AspGlyGlyValLysValThrValThrGlnAspSerSerLeuGlnAspGlyGlyPheIleTyr 120
Db 393 GACGCTGGCGCTGTACTGTAAACCAGGATTCAGTTTGCAGATGGCTGTTCATCTAC 452
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db 453 AAGTCGAAGTCATTCGCGCTGAACCTTCCTCCATGAGCACTGTATGCAAAAGAAACA 512
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLysLysGlyGly 160
Db 513 ATGGGCTGGGAAGCAGACAGCACTGACGCTTGTATCTCTGTATGGCGGTGTTGAAGAGAG 572
QY 161 IleHisLysValAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db 573 ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAAGTTGAATCAAAAGATTT 632
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db 633 TACATGGCAAGAAAGCCCTGTGACGTACCAAGGTAACCTAATGTTGACTCCAAACTGAT 692
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluAspThrGluGlyArg 220

```

```

Db 693 ATAAACAAGCCACAAGCAAGACTATACATGCTTGAGAGATGTAAGAAAGACGAGGACGC 752
QY 221 HisHisLeuPheLeu 225
Db 753 CACCATCTGTTCCTT 767
RESULT 65
AAAD03615
ID AAAD03615 standard; cDNA; 898 BP.
XX
AC AAAD03615;
XX
DT 11-SEP-2003 (revised)
DT 19-JUN-2001 (first entry)
XX
DE Dicosoma sp. red chromo/fluorescent protein, drFP583 (NFP-6) cDNA.
XX
KW Anthozoa; Chromoprotein; fluorescent protein; drFP583; NFP-6; sunscreen;
KW analyte detection assay; selectable marker; recombinant DNA application;
KW biosensor; pH indicator; invivo marker; selective filter; ss.
XX
OS Dicosoma sp.; red.
XX
FH Key Location/Qualifiers
FT CDS 93..770
FT /*tag= a
FT /product= "Chromo/fluorescent protein, drFP583 (NFP-6)"
FT /transl_except= (pos:340..600, aa:Leu-Gly)
FT /note= "insertion of 60 bases alters the reading frame"
XX
XX MO200127150-A2.
XX
PD 19-APR-2001.
XX
PF 13-OCT-2000; 2000MO-US028477.
XX
PR 14-OCT-1999; 99US-00418529.
PR 15-OCT-1999; 99US-00418917.
PR 15-OCT-1999; 99US-00418922.
PR 19-NOV-1999; 99US-00444338.
PR 19-NOV-1999; 99US-00444341.
PR 09-DEC-1999; 99US-00457556.
PR 09-DEC-1999; 99US-00457898.
PR 09-DEC-1999; 99US-00458144.
PR 09-DEC-1999; 99US-00458477.
PR 10-DEC-1999; 99MO-US029405.
PR 14-JUN-2000; 2000US-0211607P.
PR 14-JUN-2000; 2000US-0211609P.
PR 14-JUN-2000; 2000US-0211626P.
PR 14-JUN-2000; 2000US-0211627P.
PR 14-JUN-2000; 2000US-0211687P.
PR 14-JUN-2000; 2000US-0211766P.
PR 14-JUN-2000; 2000US-0211808P.
PR 14-JUN-2000; 2000US-0211888P.
PR 14-JUN-2000; 2000US-0212070P.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Terakikh A;
XX WPI; 2001-266409/27.
XX
XX P-PSDB; AAE00375.
XX
XX An Anthozoa chromo- or fluorescent protein (P1) present in an environment
XX other than its natural environment, useful as a label in analyte
XX detection assays and as a selectable marker in recombinant DNA
XX applications.
XX
XX Claim 13; Fig 6; 69pp; English.
XX
XX The present sequence is a Dicosoma sp. red chromo/fluorescent protein,
XX drFP583 (NFP-6) cDNA. NFP-6 is present in an environment other than its
XX natural environment and has an absorbance maximum ranging from 250nm to

```

CC 750nm and more usually from 540nm to 580nm and emission maximum ranging  
 CC from 275nm to 775nm and more usually from 565 to 605nm. The  
 CC chromoproteins or fluorescent proteins are useful as labels in analyte  
 CC detection assays, as selectable markers in recombinant DNA applications,  
 CC as biosensors in prokaryotic and eukaryotic cells e.g. as pH indicator  
 CC and as *in vivo* markers in animals. They are also useful in screens and  
 CC as selective filters. (Updated on 11-Sep-2003 to standardise OS field)  
 XX

Sequence 898 BP; 273 A; 173 C; 240 G; 212 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	1.78e-136	Length:	898
Score:	1184.00	Matches:	219
Percent Similarity:	98.22%	Conservative:	2
Best Local Similarity:	97.33%	Mismatches:	4
Query Match:	97.53%	Indels:	0
DB:	4	Gaps:	0

US-10-006-922a-12 (1-225) x AAD03615 (1-898)

```

OY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluLy 20
DB 93 ATGAGCTCTTCCAGAAATGTTATCAAGAGTTTCATGAGTTTAAGGTTTCGATGGAAGA 152
OY 21 ThrValaAngLyHLeGluPheGluIleGluLyGluLyGluLyValArgProTyrGluLy 40
DB 153 ACGGTCATGGGCAAGCAGATTGAAATAGAAAGGCAAGGAGGAGGAGCCATACCAAGGC 212
OY 41 HisAenThrValLyLeuLyValThrLyGlyGlyProLeuProPhaAlaTyrAspIle 60
DB 213 CACATATCCGTAACCTTAAGGTAAACCAAGGGGAGCCTTTGCCATTTGCTGGAGATTT 272
OY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80
DB 273 TTGTACCAACATTTCACTATGGAAGCAAGGTAATGTCTACAGCACCTGCCACATACCA 332
OY 81 AspTyrLyLyLeuSerPheProGluLyPheLySerTyrGluValMetAsnPhaGlu 100
DB 333 GACGTAAAGCTGTCATTTCTGAAAGATTAAATGGAAAGGTCATCAATCTTGA 392
OY 101 AspGlyLyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB 393 GACGGTGGGCGTGTACTGTAAACCAAGATTCAGTTGCAAGATGGCTGTTCACTAC 452
OY 121 LyValLyPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySerThr 140
DB 453 AAGTCAAGTTCAATGGCGTTGAACCTTCTCCGATGACCTGTTATGCAAAAGAGCA 512
OY 141 MetGlyTyrGluValSerThrGluArgLeuTyrProArgAspGlyValLeuLyGlyGlu 160
DB 513 ATGGGCTGGGAAGCCAGCACTAGGGTTGTATCTCGATGCGCGTGTGAAGAGAG 572
OY 161 IleHsLyValLeuLyLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIle 180
DB 573 ATTCAATAAGGCTCTGAAAGCTGAAGACGGTGGTCACTTACCTAGTTGAATTCAAAGTAT 632
OY 181 TyrMetAlaLyValLyProValGlnLeuProGlyTyrTyrTyrValAspSerLyLeuAsp 200
DB 633 TACATAGGCAAGAAAGCTGTGAGCTTACCAAGGTTACTTACTTATGTTGACTCCAACTGGAT 692
OY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluLyArg 220
DB 693 ATTAACAAGCACAAGAAAGCTATACATCGTTGACAGATGAAGAAACGAGGAGCCG 752
OY 221 HisHsLyLeuPheLeu 225
DB 753 CACCATCTGTCTCTT 767

```

RESULT 66  
 ADC24128  
 ID ADC24128 standard, DNA; 678 BP.  
 XX  
 AC ADC24128;

XX 18-DEC-2003 (first entry)  
 DT  
 XX  
 XX Discosoma red fluorescent protein variant fast T1 DNA.  
 DE  
 XX  
 XX Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
 KW fluorescent protein variant; transfection induction detection;  
 KW fluorescence energy resonance transfer; FRRT; protein kinase;  
 KW protein phosphatase; ion indicator; ds; mutant; fast T1.  
 KW  
 XX  
 XX Synthetic.  
 OS  
 OS Discosoma.  
 XX  
 XX US2003059835-A1.  
 EN  
 XX  
 XX 27-MAR-2003.  
 PD  
 XX  
 XX 10-APR-2002; 2002US-00121258.  
 PE  
 XX  
 XX 26-FEB-2001; 2001US-00794308.  
 PR  
 XX 24-MAY-2001; 2001US-0086538.  
 XX  
 XX (TSIE/) TSIE R Y.  
 PA (CAMP/) CAMPBELL R E.  
 XX  
 XX Tsien RY, Campbell RE;  
 PI  
 XX  
 XX WPI; 2003-743764/70.  
 DR P-PSDB; ADC24129.  
 XX  
 XX

Novel polynucleotide sequence encoding Discosoma red fluorescent protein variant having a reduced propensity to oligomerize, useful for detecting transcriptional activity.

Example 1; SEQ ID NO 3; 67bp; English.

The invention describes a polynucleotide sequence (I) encoding a Discosoma red fluorescent protein (DsRed) variant having a reduced propensity to oligomerize, comprising amino acid substitutions at the AB and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225 amino acids, given in the specification, where the substitutions result in reduced propensity of the DsRed variant to form tetramers. (I) is useful for detecting transcriptional activity by providing a host cell containing a vector which comprises (I) operatively linked to an expression control sequence, and an unit to assay the variant fluorescent protein fluorescence, and assaying fluorescence of the variant fluorescent protein produced by (VII), where variant fluorescent protein fluorescence is indicative of transcriptional activity. A polynucleotide encoding a fusion protein is useful for the analysis of *in vivo* localization or trafficking of a polypeptide of interest. A polypeptide marker is useful as markers to identify the location and amount of a target protein produced, where the target protein is fused to the marker, as a complement to or alternative for the green fluorescent protein or its spectral variant, for detecting induction of transcriptions, in applications involving fluorescence energy resonance transfer (FRET), which detects events as the function of the movement of fluorescent CC donors and acceptors towards or away from each other, for making CC fluorescent sensors for protein kinase and phosphatase activities or CC indicators for ions and molecules such as Ca<sup>2+</sup>, Zn<sup>2+</sup>, for identifying the presence of a molecule in a sample, for identifying a specific CC interaction of a first and second molecule, for determining whether a CC sample contains an enzyme or for determining the pH of the sample. (I) is useful for identifying a region or condition that regulates the activity CC of an expression control sequence. This sequence encodes Discosoma red CC fluorescent protein variant fast T1.

Sequence 678 BP; 141 A; 225 C; 205 G; 107 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	1.13e-133	Length:	678
Score:	1160.00	Matches:	216
Percent Similarity:	97.33%	Conservative:	3
Best Local Similarity:	96.00%	Mismatches:	6

Query Match: 95.55% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-006-922a-12 (1-225) x ADC24128 (1-678)

QY 1 MetArgSerSerLybAsnValIleLybGluPhenMetArgPheLybValArgMetGluGly 20  
 |||:::|||||  
 DB 1 ATGGCCCTCTCCGAGGACGTGATCAAGAGTTCATGCGCTTCAAGTCCGATGGAAGGC 60  
 QY 21 ThrValAsnGlyHsiGluPheGluIleGluGlyGluGlyGluGlyArgProTyrgLugly 40  
 ::::|  
 DB 61 TCCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCCGCCCTCAAGAGGGC 120  
 QY 41 HlsAsnThrVallybLeuLybValThrLybGlyLybProLeuProPheIleATrPApIle 60  
 |||  
 DB 121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTGGCTGGGACATC 180  
 QY 61 LeuSerProGlnPheGlnTyrgLybSerLybValTyrgValLybHsiProIleAAspIlePro 80  
 |||  
 DB 181 CTGTCCCCCAGTTCCAGTACGGCTCCAGGTGACGTGAAGCAACCCCGCGACATCCCC 240  
 QY 81 AspTyrgLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100  
 |||  
 DB 241 GACTACAAAGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCGAG 300  
 QY 101 AspGlyGlyValValThrValThrGlnAAspSerSerLeuGlnAAspGlyCybPheIleTyrg 120  
 |||  
 DB 301 GACGGCGCGGTGGTGAACCGTGAACCAAGATCTCCCTCGACGAGCGGCTCTTCATCTAC 360  
 QY 121 LybValLybPheIleGlyValAsnPheProSerAAspGlyProValMetGlnLybLybThr 140  
 |||  
 DB 361 AAGGTGAAGTTCACTGGCGTGAATCTCCCTCCACGGCCCGCTGAATGCAAGAAAGACT 420  
 QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrgProArgAAspGlyValLeuLybGlyGlu 160  
 |||  
 DB 421 ATGGCGTGGAGGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGAAGGGCGAG 480  
 QY 161 IleHsiLybAlaLeuLybLeuLybAAspGlyGlyHsiTyrgLeuValGluPheLybSerIle 180  
 |||  
 DB 481 ATCCACAAAGGCCCTGAAGCTGAAGACGGCGGCCACTACCTGTGGAGTTCAAGTCCATC 540  
 QY 181 TyrgMetAlaLybLybProValGlnLeuProGlyTyrgTyrgTyrgValAAspSerLybLeuAAsp 200  
 |||  
 DB 541 TACATGGCCAAAGAACCGGTGACGTGCGCGCTCACTACCTACGTGACTCCAACTGGAC 600  
 QY 201 IleThrSerHsiAAsnGluAAspTyrgThrIleValGluGlnTyrgIuArgThrGluGlyArg 220  
 |||  
 DB 601 ATCACTCCCAACAAGAGACTACACCATCGTGAAGAGTACGAGCGGCGGAGGGCGCG 660  
 QY 221 HlsHsiLeuPheLeu 225  
 DB 661 CACCACTGTCTCTG 675

RESULT 67  
 ADL46207 standard; DNA; 678 BP.  
 ID ADL46207  
 AC ADL46207;  
 XX 20-MAY-2004 (first entry)  
 DT XX  
 DE Discosoma red fluorescent protein (DsrEd) variant T1 coding sequence.  
 KM de; gene; red fluorescent protein; DsrEd; fluorescence; red wavelength;  
 XX oligomerization; tetramerization; immunoassay; hybridization assay.  
 OS Discosoma sp.  
 FH Key Location/Qualifiers  
 FT 1..678  
 FT CDS /\*tag= a  
 FT product= "DsrEd variant T1 protein"

PN WO2003086446-A1.  
 XX  
 XX 23-OCT-2003.  
 XX  
 PF 09-APR-2003; 2003WO-US010879.  
 XX  
 PR 10-APR-2002; 2002US-00121258.  
 XX 29-JUL-2002; 2002US-00209208.  
 XX  
 PA (RBGC ) UNIV CALIFORNIA.  
 XX  
 PI Tsien RY, Campbell RE, Baird GS;  
 XX  
 DR WPI; 2003-845265/78.  
 DR P-PSDB; ADL46206.  
 XX  
 PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
 PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
 PT in molecular biology, e.g. in immunoassays or in tracking protein  
 PT movement in cells.  
 XX  
 PS Disclosure; SEQ ID NO 5; 166pp; English.  
 XX  
 CC The invention relates to a polynucleotide sequence encoding a Discosoma  
 CC red fluorescent protein (DsrEd) variant having a reduced propensity to  
 CC oligomerize. The protein variant comprises one or more amino acid  
 CC substitutions at the AB and/or AC interface(s) of the wild-type DsrEd  
 CC sequence, where the substitutions result in reduced propensity of the  
 CC DsrEd variant to form tetramers and where the variant displays detectable  
 CC fluorescence of at least one red wavelength. The composition and methods  
 CC are useful in producing red fluorescent proteins having reduced  
 CC propensity for oligomerization, especially tetramerization. The protein  
 CC may be used in molecular biology and in other scientific applications,  
 CC such as in immunoassays or hybridization assays, or in tracking the  
 CC movement of proteins in cells. This sequence corresponds to the DsrEd  
 CC coding sequence for the variant T1.  
 XX

SEQ Sequence 678 BP; 141 A; 225 C; 205 G; 107 T; 0 U; 0 Other;  
 SQ

Alignment Scores:  
 Pred. No.: 1,13e-133 Length: 678  
 Score: 1160.00 Matches: 216  
 Percent Similarity: 97.33% Conservative: 3  
 Best Local Similarity: 96.00% Mismatches: 6  
 Query Match: 95.55% Indels: 0  
 DB: 11 Gaps: 0

US-10-006-922a-12 (1-225) x ADL46207 (1-678)

QY 1 MetArgSerSerLybAsnValIleLybGluPhenMetArgPheLybValArgMetGluGly 20  
 |||  
 DB 1 ATGGCCCTCTCCGAGGACGTGATCAAGAGTTCATGCGCTTCAAGTCCGATGGAAGGC 60  
 QY 21 ThrValAsnGlyHsiGluPheGluIleGluGlyGluGlyGluGlyArgProTyrgLugly 40  
 ::::|  
 DB 61 TCCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCCGCCCTCAAGAGGGC 120  
 QY 41 HlsAsnThrVallybLeuLybValThrLybGlyLybProLeuProPheIleATrPApIle 60  
 |||  
 DB 121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTGGCTGGGACATC 180  
 QY 61 LeuSerProGlnPheGlnTyrgLybSerLybValTyrgValLybHsiProIleAAspIlePro 80  
 |||  
 DB 181 CTGTCCCCCAGTTCCAGTACGGCTCCAGGTGACGTGAAGCAACCCCGCGACATCCCC 240  
 QY 81 AspTyrgLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100  
 |||  
 DB 241 GACTACAAAGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCGAG 300  
 QY 101 AspGlyGlyValValThrValThrGlnAAspSerSerLeuGlnAAspGlyCybPheIleTyrg 120  
 |||  
 DB 301 GACGGCGCGGTGGTGAACCGTGAACCAAGATCTCCCTCGACGAGCGGCTCTTCATCTAC 360

```
QY 121 LysValIysPheIleGlyValAsnPheProSerAaspGlyProValMetGlnLysLysThr 140
DB 361 AAGGTGAAGTTCAATCGCGCGTGAACCTCCCTCCGACGGCCCGCTAAATGCAAGAAAGACT 420
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAaspGlyValIleuLysGlyGlu 160
DB 421 ATGGGCTGGAGGGCTCCACCGACCGCTGTACCCCGCGACGGCGTGGTGAAGGGCGAG 480
QY 161 IleHisValAlaLeuLysLeuLysAaspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB 481 ATCCACAAAGCCCTCGAAGCTTAAGACGCGCGCCACTACCTGGTGGAGTTCAAGTTCATC 540
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAaspSerLysLeuAasp 200
DB 541 TACATGGCAAGAAAGCCCGTGCAGCTGCCGCTACTACTAGAGACTCCAAAGCTGGAC 600
QY 201 IleThrSerHisAaspGluAaspTyrThrIleValGluGlnTyrGluAArgThrGluGlyArg 220
DB 601 ATCACTCTCCCAACAGAGACTACACATCGTGAAGAGTACGAGCGCGCGCGGCGCCGC 660
QY 221 HisHisLeuPheLeu 225
DB 661 CACCACTGTCTCTG 675

RESULT 68
ADN33980
ID ADN33980 standard; DNA; 704 BP.
AC ADN33980,
XX 01-JUL-2004 (first entry)
DE Mutant-type DARED encoding sequence.
XX Mutant-type DARED encoding sequence.
XX Cnidarian; fluorescence resonance energy transfer; FRET; wild-type DARED;
KM mutant; muteln; da.
XX
XX Diicosoma sp.
XX
XX WO2003054158-A2.
XX
XX 03-JUL-2003.
XX
XX 18-DEC-2002; 2002WO-US040539.
XX
XX 19-DEC-2001; 2001US-0341723P.
XX
XX (UYCH-) UNIV CHICAGO.
XX
XX Beye B, Glick B;
XX
XX WPI; 2003-569236/53.
XX
XX Claim 7, SEQ ID NO 3; 65bp; English.
XX
XX The present invention relates to nucleic acid that encodes a rapidly
XX maturing chromo- or fluorescent mutant of a Cnidarian chromo- or
XX fluorescent protein or its mutant. The protein is useful in applications
XX involving nucleic acid encoding a chromo- or fluorescent protein and is
XX useful for producing a chromo- and/or fluorescent protein which involves
XX growing the cell, whereby the protein is expressed, and isolating the
XX protein substantially free of other proteins. The protein is useful in
XX applications involving chromo- or fluorescent protein and is useful as
XX PCR primers, hybridization probes, etc. The expression cassettes are
XX useful for synthesizing related proteins. The chromoproteins are useful
XX as coloring agents which are capable of imparting color or pigment to a
XX particular composition of matter e.g. food compositions, pharmaceuticals,
XX cosmetics, living organisms, e.g., animals and plants. The chromoproteins
XX may also find use as labels in analyte detection assays, e.g. assays for
```

```
CC biological analytes of interest and as selectable markers in recombinant
CC DNA applications, e.g. the production of transgenic cells and organisms.
CC The fluorescent proteins find use in a variety of different applications,
CC e.g. in fluorescence resonance energy transfer (FRET) applications, as
CC biosensors in prokaryotic and eukaryotic cells, in applications involving
CC the automated screening of arrays of cells expressing fluorescent
CC reporting groups by using microscopic imaging and electronic analysis, as
CC second messenger detectors, and in fluorescence activated cell sorting
CC applications and as in vivo marker in animals. The fluorescent proteins
CC also find use in protease cleavage assays. The proteins can also be used
CC in assays to determine the phospholipid composition in biological
CC membranes and as a fluorescent timer. The present sequence represents the
CC mutant-type DARED encoding sequence.
XX
XX Sequence 704 BP; 145 A; 237 C; 212 G; 110 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 1,19e-133 Length: 704
Score: 1160.00 Matches: 216
Percent Similarity: 97.33% Conservative: 3
Best Local Similarity: 96.00% Mismatches: 6
Query Match: 95.55% Indels: 0
DB: Gaps: 0

US-10-006-922A-12 (1-225) x ADN33980 (1-704)
QY 1 MetArgSerSerLysAasnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
DB 20 ATGGCTCTCTCCGAGAGAGCTCATCAAGAGATTCAATGCCCTTCAAGATGCCATGAGAGGC 79
QY 21 ThrValAaspGlyHisGluPheGluIleGluGlyGlyGlyArgProTyrGluGly 40
DB 80 TCCGTGAACGGGCAAGAGTTGAGATTCAGAGGCGAGGCGCGCGCCCTTACGAGGGC 139
QY 41 HisAasnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAaspIle 60
DB 140 ACCCAGACCGCCAAAGCTGAAGTGAACCAAGGGCGGCCCTGCCCTTGGCGGAGCATTC 199
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAaspIlePro 80
DB 200 CTGTCCCGCCAGATTCCAGTACGAGTCCCAAGGTTGATCGTGAAGCAACCCCGCAGCATCCCC 259
QY 81 AaspTyrLysLysLeuSerPhePheProGlnLysPheLysTrpGluAArgValMetAaspPheGlu 100
DB 260 GACTACAAAGAAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGGCTGATGAACCTTCAG 319
QY 101 AaspGlyValAlaThrValThrGlnAaspSerSerLeuGlnAaspGlyCysApheIleTyr 120
DB 320 GACGGCGCGGTGTGACCGGTGACCGGAGACTCTCTCGAGAGAGCGCTCTTATCTAC 379
QY 121 LysValIysPheIleGlyValAsnPheProSerAaspGlyProValMetGlnLysLysThr 140
DB 380 AAGGTGAAGTTCAATCGCGGTGAACCTCCCTCCGACGGCCCGTAAATGCAAGAAAGACT 439
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAaspGlyValIleuLysGlyGlu 160
DB 440 ATGGGCTGGAGGGCTCCACCGACGGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 499
QY 161 IleHisValAlaLeuLysLeuLysAaspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB 500 ATCCACAAAGCCCTTAAGCTTAAGACGCGCGCCACTACCTGGTGAAGTTCAAGTTCATC 559
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAaspSerLysLeuAasp 200
DB 560 TACATGGCAAGAAAGCCCGTGCAGCTGCCGCTACTACTAGAGACTCCAAAGCTGGAC 619
QY 201 IleThrSerHisAaspGluAaspTyrThrIleValGluGlnTyrGluAArgThrGluGlyArg 220
DB 620 ATCACTCTCCCAACAGAGACTACACATCGTGAAGAGTACGAGCGCGCGCGGCGCCGC 679
QY 221 HisHisLeuPheLeu 225
DB 680 CACCACTGTCTCTG 694
```

RESULT 69  
ADM13536  
ADM13536 standard; DNA; 678 BP.  
XX  
XX ADM13536;  
AC  
AD  
XX 15-JUL-2004 (first entry)  
XX  
XX DNA encoding superfolding coral DsRed protein #1.  
DE  
XX  
XX GFP; superfolding mutant; ds; green fluorescent protein;  
KM directed-evolution; folding interference domain; reporter protein;  
KW mutant; DsRed; coral.  
XX  
XX Diacoboma sp.  
OS Synthetic.  
XX  
XX US2004078148-A1.  
PN  
XX 22-APR-2004.  
PD  
XX 24-APR-2003; 2003US-00423688.  
PF  
XX 24-APR-2002; 2002US-00132067.  
PR  
XX (WALD/) WALDO G. S.  
PA  
XX  
XX Waldo GS;  
PI  
XX WPI; 2004-340059/31.  
DR P-PSDB; ADM13537.  
XX  
XX New modified green fluorescent protein, useful as a reporter in expression studies.  
PT  
XX  
XX Disclosure; SEQ ID NO 40; 46pp; English.  
PS  
XX The invention relates to a green fluorescent protein (GFP) that comprises at least 80% identity to the GFP superfolding mutant appearing as CC  
XX ADM13535 that has at least one amino acid substitution selected from CC  
XX e.g., a substitution at position 30 that is an arginine or a conservative CC  
XX variant of arginine and measurable fluorescence activity. Also included CC  
XX are directed-evolution for generating an enhanced folding variant of a CC  
XX target polypeptide (comprising mutating a polynucleotide encoding CC  
XX polypeptide of interest to generate a library of mutated polynucleotides, CC  
XX linking mutated polynucleotides to polynucleotide encoding folding CC  
XX interference domain to form fusion protein (FP) constructs, expressing CC  
XX FP, and selecting FP that display optimal folding activity in relation to CC  
XX FP comprising wild-type polypeptide and folding interference domain, thus CC  
XX identifying polypeptide with enhanced folding activity), an enhanced CC  
XX folding variant of a fluorescent protein generated by the method, an CC  
XX enhanced folding variant of a chromophoric protein generated by the CC  
XX method, enhancing folding of a polypeptide comprising multiple domains CC  
XX (involving joining a first domain of the polypeptide to a poorly folding CC  
XX domain, to form a fusion protein, mutating the first domain, detecting an CC  
XX increase in the amount of activity generated by a first mutated fusion CC  
XX protein in comparison to a fusion protein comprising a wild-type first CC  
XX domain and the poorly folding polypeptide domain, thus identifying a CC  
XX first domain with enhanced folding, joining a second domain of the CC  
XX polypeptide to the first mutated fusion protein to form a second fusion CC  
XX protein, mutating the second domain, and detecting an increase in the CC  
XX amount of activity generated by a second mutated fusion protein in CC  
XX comparison to a fusion protein comprising the wild-type second domain and CC  
XX the first mutated fusion protein, thus identifying a target polypeptide CC  
XX with multiple domains that have enhanced folding), an isolated nucleic CC  
XX acid encoding the GFP, an expression vector comprising the nucleic acid CC  
XX and a host cell comprising the vector. The DsRed protein from coral is CC  
XX also considered for mutation using the above method. The enhanced folding CC  
XX variants are useful as reporter proteins to express the report level of a CC  
XX protein. The method is useful for directed-evolution of generating CC  
XX enhanced folding variant of target polypeptide and for improving folding CC  
XX and solubility of a target protein. The present sequence encodes a coral

CC DsRed enhanced folding variant. NOTE: The authors refer to SEQ ID 38-42  
CC as SEQ ID 1-5 throughout the specification and never refer to SEQ ID 1-37  
CC by number, it is obvious that SEQ ID 38-42 are meant to be SEQ ID 1-5.  
XX  
SQ Sequence 678 BP; 203 A; 127 C; 181 G; 167 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3,32e-130 Length: 678  
Score: 1132.00 Matches: 208  
Percent Similarity: 95.56% Conservative: 7  
Best Local Similarity: 92.44% Mismatches: 10  
Query Match: 93.25% Indels: 0  
DB: Gaps: 0  
US-10-006-922a-12 (1-225) x ADM13536 (1-678)  
QY 1 MetArgSerSerIysAsnValIleLySGluPheMetArgPheIysValIArgMetGluGly 20  
DB 1 ATGGAAGCTTCCGAGGATGTTATCAAGAGCTTCATGAGTTTAAAGTTCAATGAGGA 60  
QY 21 ThrValAsnGlyIhIeGluPheGluIleGluGluGluGluGluGluGluGluGluGluGlu 40  
DB 61 TCGGTCAATGGGACGAGTTGAAATAGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
QY 41 HisAsnThrValIysIleuIysValIThrLySGlyIProLeuProPheIaITPAspIle 60  
DB 121 ACCCAGAACGTAAAGCTTAAGTACTTAAGGGGGGACCTTGCATTGCTGGGATATT 180  
QY 61 LeuSerProGluPheGluIysGlySerIysValIYrValIysHisPhePheIlePro 80  
DB 181 TTGTCAACCAATTCAGTATGAGAGCAGATGATATGCAAGACCCCTGCCGACATACCA 240  
QY 81 AspTyrIleIysLeuSerPheProGluGlyPheLySGluIysValIleuIysValIleuIys 100  
DB 241 GACTATTAAGGCTGCTCATTTCTGAAAGATTTAAATGGAGAGGCTCATGAACTTGAA 300  
QY 101 AspGlyGlyValValIThrValIThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCTTCGCTACTGTAACCCAGATTCAGATTGGAGAGATGGCTGTGATCTTAC 360  
QY 121 IysValIysPheIleGlyValIAsnIlePheProSerAspGlyProValMetGlnIysLeuThr 140  
DB 361 AAGGTCAAGTTCAATGGCGGAACTTCTTCGATGACCTGTTATGCAAAAGAAACAA 420  
QY 141 MetGlyTTPGluIAsnThrGluIysGluIysPheProArgAspGlyValIleuIysGlyGlu 160  
DB 421 ATGGGCTGGGAAACCGAACCTGAGCGTTGTATCTCTGTATGCGCTGTGAAAGAGAT 480  
QY 161 IleHisIysIleuIysLeuIysAspGlyGlyHisTyrIleuValIleuIysSerIle 180  
DB 481 ATTCAATAGGCTCTGAAAGCTGAAAGACGGTGTCTTACTTACTGATGATCAAAAGTATT 540  
QY 181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrValIAspSerIysLeuAsp 200  
DB 541 TACATGGCAAAAGAGCTGACAGCTACAGGGGACTACTATGTTGACTCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValIleGluIysGluIysGluIysGluIys 220  
DB 601 ATTAACAACCAACCAAGAAAGCTATCAATGTTGAGCATATGAAAGACCGAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCTT 675  
RESULT 70  
ADC24130  
ID ADC24130 standard; DNA; 681 BP.  
XX  
XX ADC24130;  
AC  
XX 18-DEC-2003 (first entry)  
DT  
XX  
XX Diacoboma red fluorescent protein variant dimer2 DNA.



XX Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
 KW fluorescent protein variant; transcription induction detection;  
 KM fluorescence energy resonance transfer; FRRT; protein kinase;  
 KM protein phosphatase; ion indicator; mutant; dimer2; ds.  
 OS Synthetic.  
 OS Discosoma.  
 XX US2003059835-A1.  
 PN 27-MAR-2003.  
 PD 10-APR-2002; 2002US-00121258.  
 PF 26-FEB-2001; 2001US-00794308.  
 PR 24-MAY-2001; 2001US-0086538.  
 XX (TSIE/) TSIE R Y.  
 PA (CAMP/) CAMPBELL R E.  
 XX Tsien RY, Campbell RB;  
 PI WPI; 2003-743764/70.  
 DR P-PSDB; ADC24131.  
 XX Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
 PT variant having a reduced propensity to oligomerize, useful for detecting  
 PT transcriptional activity.  
 PS Example 1; SEQ ID NO 5; 67pp; English.  
 XX The invention describes a polynucleotide sequence (I) encoding a  
 CC Discosoma red fluorescent protein (DsRed) variant having a reduced  
 CC propensity to oligomerize, comprising amino acid substitutions at the AB  
 CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
 CC amino acids, given in the specification, where the substitutions result  
 CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
 CC useful for detecting transcriptional activity by providing a host cells  
 CC containing a vector which comprises (I) operatively linked to an  
 CC expression control sequence, and an unit to assay the variant fluorescent  
 CC protein fluorescence, and assaying fluorescence of the variant  
 CC fluorescent protein produced by (VII), where variant fluorescent protein  
 CC fluorescence is indicative of transcriptional activity. A polynucleotide  
 CC encoding a fusion protein is useful for the analysis of in vivo  
 CC localization or trafficking of a polypeptide of interest. A polypeptide  
 CC marker is useful as markers to identify the location and amount of a  
 CC target protein produced, where the target protein is fused to the marker,  
 CC as a complement to or alternative for the green fluorescent protein or  
 CC its spectral variant, for detecting induction of transcription, in  
 CC applications involving fluorescence energy resonance transfer (FRET),  
 CC which detects events as the function of the movement of fluorescent  
 CC donors and acceptors towards or away from each other, for making  
 CC fluorescent sensors for protein kinase and phosphatase activities or  
 CC indicators for ions and molecules such as  $Ca^{2+}$ ,  $Zn^{2+}$ , for identifying the  
 CC presence of a molecule in a sample, for identifying a specific  
 CC interaction of a first and second molecule, for determining whether a  
 CC sample contains an enzyme or for determining the pH of the sample. (I) is  
 CC useful for identifying a region or condition that regulates the activity  
 CC of an expression control sequence. This sequence encodes Discosoma red  
 CC fluorescent protein variant dimer2.

SO Sequence 681 BP; 145 A; 231 C; 205 G; 100 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 7.71e-129 Length: 681  
 Score: 1121.00 Matches: 207  
 Percent Similarity: 95.96% Conservative: 7  
 Best Local Similarity: 92.83% Mismatches: 9  
 Query Match: 92.34% Indels: 0  
 DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x ADC24130 (1-681)

QY 3 SerSerIysAsnValIleIyGluPheMetArgPheIysValArgMetGluIThrVal 22  
 DB 10 TCCCTCCAGGACGTCACCAAGAGTTTCATGCTTCAAGAGGCGCATGAGAGGCTCCGTG 69  
 QY 23 AsnGlyHISGluPheGluIleGluGluIyGluIyValArgProIyGluIyHISAsn 42  
 DB 70 AACGGCCACGAGTTCAGATCTCAGAGGCGAGGCGAGGCGCCCTTCAGAGGCAACCAG 129  
 QY 43 ThrValIysLeuIyValIThrIyGluIyGluIyProLeuProPheIATPASPILeLeuSer 62  
 DB 130 ACCGCCAAGCTGAGGATCGACCAAGGCGGCGCCCTGCGCTTCGCTCGAGCATCTGCTCC 189  
 QY 63 ProGluPheGluIyGlySerIyValIyValIyHISProAlaAPILeProASPlyr 82  
 DB 190 CCCCAAGTTCAGTTCAGGCTCCAGGCGGTACCTGAGACACCCCGGACATCCCGACTAC 249  
 QY 83 LysIyLeuSerPheProGluGluIyPheIySTrpgIuArgValIMetAsnPheGluASPly 102  
 DB 250 AAGAAGCTGTCTTCCCGAGGCTTCAGTGGAGCCCGTGAATGAATCTGAGGACGGC 309  
 QY 103 GlyValIyThrValIThrIyAspSerSerIeGluIyAspGlyCySPheIleTyIyVal 122  
 DB 310 GCGGTGTGACCGTGAACCAAGACTCCCTGACAGACGCGACGCTGATCTACAGGTG 369  
 QY 123 LysPheIleGlyValIAsnPheProSerAspGlyProValMetGluIyIyThrMetGly 142  
 DB 370 AAGTTCGCGGACCAACTTCCCGCGCGCCCGCTGAAGAGAAACACATGGGCGC 429  
 QY 143 TrpGluIAserThrIyArgIyLeuTyProArgAspGlyValIleuIyGluIyIleHis 162  
 DB 430 TGGAGGCTTCACCGAGCGCTGTACCCCGACAGCGCGTGTGAGAGGAGATCCAC 489  
 QY 163 LysAlaIleuIyLeuIyAspGlyGlyIyHISTyIeValIyGluPheIySerIleTyIyMet 182  
 DB 490 CAGCGCTCGAAGCTGAGAGACGAGGCGCACTACGTGGTGAATCAAGACCATCTACATG 549  
 QY 183 AlaIyIyProValIyGluIyLeuProGlyTyTyTyTyValIAspSerIyLeuAspIleThr 202  
 DB 550 GCCAAGAAGCCCGTGCACCTGCCGCTACTACAGTGAACACMACACTGACATGCACC 609  
 QY 203 SerHISAsnGluAspTyThrIleValIyGluIyArgIyThrGluIyArgHISHis 222  
 DB 610 TCCCAACAGAGACTACACATCTGGAACAGTACGCGCTCCGAGGCGCGCCACAC 669  
 QY 223 LeuPheIy 225  
 DB 670 CTGTTCTCTG 678  
 DB ADL46209  
 ID ADL46209 standard; DNA; 681 BP.  
 XX ADL46209;  
 AC ADL46209;  
 XX 20-MAY-2004 (first entry)  
 DT 20-MAY-2004 (first entry)  
 XX Discosoma red fluorescent protein (DsRed) variant dimer2 coding sequence.  
 DE ds; gene; red fluorescent protein; DsRed; fluorescence; red wavelength;  
 KW oligomerization; tetramerization; immunosassay; hybridization assay.  
 XX Discosoma sp.  
 OS Discosoma sp.  
 XX Key Location/Qualifiers  
 FT CDS 1..681  
 FT /tag= a  
 FT /product= "DsRed variant dimer2 protein"  
 XX WO2003086446-A1.  
 XX 23-OCT-2003.

PF 09-APR-2003; 2003WO-US010879.  
 XX  
 PR 10-APR-2002; 2002US-00121258.  
 PR 29-JUL-2002; 2002US-00209208.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Tsien RY, Campbell RE, Baird GS;  
 DR WPI; 2003-945265/78.  
 DR P-PSDB; ADL46208.  
 XX  
 PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
 PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
 PT in molecular biology, e.g. in immunoassays or in tracking protein  
 PT movement in cells.  
 XX  
 PS  
 XX Disclosure; SEQ ID NO 7; 166pp; English.  
 CC The invention relates to a polynucleotide sequence encoding a Discosoma  
 CC red fluorescent protein (DRed) variant having a reduced propensity to  
 CC oligomerize. The protein variant comprises one or more amino acid  
 CC substitutions at the AB and/or AC interface(s) of the wild-type DRed  
 CC sequence, where the substitutions result in reduced propensity of the  
 CC DRed variant to form tetramers and where the variant displays detectable  
 CC fluorescence of at least one red wavelength. The composition and methods  
 CC are useful in producing red fluorescent proteins having reduced  
 CC propensity for oligomerization, especially tetramerization. The protein  
 CC may be used in molecular biology and in other scientific applications,  
 CC such as in immunoassays or hybridization assays, or in tracking the  
 CC movement of proteins in cells. This sequence corresponds to the DRed  
 CC coding sequence for the variant dimer2.  
 XX  
 SQ Sequence 681 BP; 145 A; 231 C; 205 G; 100 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 7,71e-129 Length: 681  
 Score: 1121.00 Matches: 207  
 Percent Similarity: 95.96% Conservative: 7  
 Best Local Similarity: 92.83% Mismatches: 9  
 Query Match: 92.34% Indels: 0  
 DB: Gaps: 0  
 US-10-006-922A-12 (1-225) x ADL46209 (1-681)  
 QY 3 SerSerTybAenValIleVbGluPhMeKArGphelysValaArgMeGluGlyThrVal 22  
 DB 10 TCCTCCGAGGACGTCACTCAAGAGTTCAATGCGCTTCAAGGTCGATGAGGGCTCCGTG 69  
 QY 23 AaNgIyHAgIuPhaGluIleGluGlyGluGlyATaTProTyTgIuGlyHAsn 42  
 DB 70 AACGGCCACAGTTTCAGATGAGGGCGAGGGCCAGGGCCCTTACGAGGGACCCAG 129  
 QY 43 ThrValIyAseuLybValThrLyAGlyGlyProLeuProPhaIaTPAaPileLeuSer 62  
 DB 130 ACCGCCAAAGCTGAGAGGTGACCAAGGGGGCCCTGCGCTTCCGTGGACATCTCTGCC 189  
 QY 63 ProGlnPhaGlnTyGlySerTybValTyValIyShiSPoAlaAaPileProAaPTyr 82  
 DB 190 CCCCAAGTTCAGTACGGCTCCAGGGCGTACGTGAAGCACCCCGCGCATCTCCGAACTAC 249  
 QY 83 LysAlLybAseuSerPheProGluGlyPheLyATrGluArgValMeLanPhaGluAaPgly 102  
 DB 250 AAGAAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGACCGCGTGAATCTTGAAGACGGC 309  
 QY 103 GlyValValThrValThrGlnAaSPSerSerLeuGlnAaPglyCybPheIleTyTlybVal 122  
 DB 310 GGGCGTGGTGAACGTGACCCAGAGACTCTCCCTGACAGGACGGCAGCGATCTTACAAGGTG 369  
 QY 123 LysPheIleGlyValaAaPhaProSerAaPglyProValMeGlnIyGlybThrMeGly 142  
 DB 370 AAGTTCCGCGGACCACTTCCCGCCCGCGCGCCCGTAAATGACAGAAAGAACCATGGGC 429

QY 143 TrpGluAlaSerThrGluArgLeuTyProAaArgAaPglyValaLeuLybGlyGluIleHis 162  
 DB 430 TGGGAGGGCTCCACCGAGGGCTGTACCCCGGAGAGGGGTGTAAGGGCGAGATCCAC 469  
 QY 163 LysAlaLeuLybAseuLybAaPglyGlyHAsTyTLeuValGluPhaLySerIleTyTMeC 182  
 DB 490 CAGGCCCTGAAGCTGAAGGACGGCGCACTACCTGTGGAGTTCAAGACATCTACATG 549  
 QY 183 AlaLybAseuProValGlnLeuProGlyTyTTyTTrValAaSPSerTybAaPileThr 202  
 DB 550 GCCAAGAGGCCCGTGAAGCTGCGCGCTTACTTACTGACACCAAGCTGGACATCAC 609  
 QY 203 SerHisAaNgIyAaPlyTyThrIleValGluGlnTyGluArgThrGluArgHisHis 222  
 DB 610 TCCCAACAGAGGACTACACCATCGTGAACAGATGAGAGGCTCCGAGGGCCCGCACAC 669  
 QY 223 LeuPheLeu 225  
 DB 670 CTGTCTCTG 678  
 RESULT 72  
 AAA48748  
 ID AAA48748 standard; cDNA; 678 BP.  
 XX  
 AC AAA48748;  
 XX 19-SEP-2000 (first entry)  
 DT  
 XX Humanised anthozoa novel fluorescent protein drFP583/dmFP592 hybrid cDNA.  
 DE  
 XX Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
 KM fluorescent labeling; hybrid; ss.  
 XX  
 OS Discosoma sp; "red".  
 OS Synthetic.  
 FH  
 FH Key Location/Qualifiers  
 FT CDS 1..678  
 XX /product= "drFP583/dmFP592 hybrid protein"  
 PN W0200034326-A1.  
 XX  
 PD 15-JUN-2000.  
 XX  
 PF 10-DEC-1999; 99MO-US029473.  
 XX  
 PR 11-DEC-1998; 98US-00210330.  
 PR 14-OCT-1999; 99US-00418529.  
 XX  
 PA (CLON-) CLONTECH LAB INC.  
 XX  
 PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
 PI Ding U;  
 XX  
 XX WPI; 2000-423381/36.  
 DR P-PSDB; AAY99838.  
 XX  
 PT Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
 PT useful for fluorescent labeling and as markers.  
 PS  
 PS Claim 6; Page 77; 86pp; English.  
 CC The present sequence is the nucleotide sequence of humanised  
 CC drFP583/dmFP592. Amplified wild-type coding region fragments from drFP583  
 CC and dmFP592 were digested with restriction endonucleases and combined  
 CC together to produce a hybrid construct. drFP583 is a full-length cDNA  
 CC encoding a novel fluorescent protein (nFP) from Discosoma sp. "red", a  
 CC non-bioluminescent species of the Class Anthozoa. The hybrid sequence was  
 CC changed to optimise the codons for expression of the protein in mammalian  
 CC cells. Fluorescent proteins can be used in fluorescent labeling, a useful  
 CC tool for marking a protein, cell or organism of interest. Unlike other  
 CC markers used in protein labeling, such as beta-galactosidase and  
 CC luciferase, fluorescent proteins do not require an exogenous cofactor or

CC substrate. Methods involving fluorescent proteins are also less laborious  
 CC and less difficult to control than the traditional methods of fluorescent  
 CC labeling, where a protein of interest is purified and then covalently  
 CC conjugated to a fluorophore derivative. Novel fluorescent proteins  
 CC isolated from species of the Class Anthozoa can be used as markers for  
 CC gene expression and protein localization studies, and in fluorescence  
 CC resonance energy transfer (FRET) reactions. They may have improved  
 CC properties and better suitability for larger excitations compared to  
 CC prior art fluorescent proteins such as green fluorescent protein

XX Sequence 678 BP; 145 A; 217 C; 218 G; 98 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	1.36e-128	Length:	678
Score:	1119.00	Matches:	206
Percent Similarity:	96.89%	Conservative:	12
Best Local Similarity:	91.56%	Mismatches:	7
Query Match:	92.17%	Indels:	0
DB:	3	Gaps:	0

US-10-006-922A-12 (1-225) x AAA48748 (1-678)

```

OY 1 MetArgSerSerLYaAnValIleYsgLuphMetAArgPheLYaIArgMetGluGly 20
DB 1 ATGAGCTGCAGCAAGAGATGATCAAGAGATTCAATCGGTTCAAGTCCGATGAGGAC 60
OY 21 ThrValaAngLYHAGLuphGluIleGluGluGluGluGluGluGluGluGluGluGlu 40
DB 61 ACCGTGAACGGCCACGAGATTTCAGATCAAGGGCGAGGGCGAGGGCGGCTTACAGAGG 120
OY 41 HIsAnthrValLYeLyuLYaValThriYsgLYiProLeuProPheAlaTriPheIle 60
DB 121 CACTGCACAGCGAAGCTCATGGTGAACCAAGGGCGGGCCCTTCCCTTCCATCACTTC 180
OY 61 LeuSerProGlnPheGlnTyrGlySerLYaValYrValLYeHIsProAlaAspIlePro 80
DB 181 CTCAGCCCCCAGTTTCAGTACGAGCAAGAGTGTACGTAAACACCCGCGACATCCCC 240
OY 81 AspTyrLYeLYeLyuSerPheProGluGlyPheLYSTPGLYaIArgValMetAnPheGlu 100
DB 241 GACTCAAGAGAGCTACGCTCCCGCAGGGCTTCAAGTGGAGGGGCTGATGAATTCGAG 300
OY 101 AspGlyLYaValValThriValThriGlnAspSerSerLeuGlnAspGlyCYaPheIleTyr 120
DB 301 GACGGCGGCGGTGTACCGTGAAGCAGACACCTTCAGAGAGCGGCTTCATCTAC 360
OY 121 LybValLYePheIleGlyLYaValAsnPheProSerAspGlyProValMetGlnLYeIsthr 140
DB 361 GAGGTGAAGTTTCATCGCGGTGAATCTCCCGCAGCGAGCGGCCGTGATGACGGCGGACC 420
OY 141 MetGlyTPGGLyALAserThrgLYaIArgLeuTyrProArgAspGlyValLeuLYeGlu 160
DB 421 CGGGGCTGGAGGGCCAGCGAGCGGCTTACCCCGGGAAGCGGCTGTCMAAGGCCAGC 480
OY 161 IleHIsLYaLYeLYeLYeLYeLYeLYeLYeLYeLYeLYeLYeLYeLYeLYeLYeLYe 180
DB 481 ATCCACATAGCGCTCCGCGTCAAGGGCGGCGGCACTACTCGTGAAGTTCAAGAGCATTC 540
OY 181 TyrMetAlaLYeLYeProValGlnLeuProGlyTyrTyrTyrValAspSerLYeLeuAsp 200
DB 541 TACATAGGCCCAAGAGCGCTGACGCTCCCGGCTACTACGTGAGCAGAGCTCGAGC 600
OY 201 IleThrSerHIsAngLYuAspTyrThriLYeValGluGluGluGluGluGluGluGlu 220
DB 601 ATCACCAGGACCAAGAGACTACACCATCTGTGAGCACTACAGAGGAGCCAGGGCGGG 660

```

OY 221 HIsHIsLeuPheLeu 225  
 DB 661 CACCACTCTCTCCTC 675

RESULT 73  
 AAD46280  
 ID AAD46280 standard; DNA; 678 BP.

XX AAD46280;  
 AC  
 XX  
 DT 27-DEC-2002 (first entry)  
 XX  
 DE Diacosoma sp. dFp583/dmFp592 (6/90) protein encoding DNA.  
 XX  
 KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
 KW fluorescence activated cell sorting application; fluorescent timer;  
 KW biosensor; fluorescence resonance energy transfer application; FRET;  
 KW colouring agent; recombinant DNA application; analyte detection assay;  
 KW sunscreen; second messenger detector; dFp583 protein; dmFp592 protein;  
 KW gene; ds.  
 XX  
 OS Diacosoma sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..678  
 FT /\*tag= a  
 FT /product= "dFp583/dmFp592 (6/90) protein"

MO200268459-A2.

06-SEP-2002.

20-FEB-2002; 2002MO-US005749.

21-FEB-2001; 2001US-0270983P.

04-DEC-2001; 2001US-00006922.

(CLON-) CLONTECH LAB INC.

Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Fradkov A;

WPI; 2002-691654/74.

P-PSDB; AAB28835.

PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
 of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
 PT analyte detection assays or fluorescence activated cell sorting  
 applications.

PS Disclosure; Page 72; 80pp; English.

XX The invention relates to nucleic acid molecules encoding non-aggregating  
 CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
 CC useful in analyte detection assays, as colouring agents, as markers in  
 CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
 CC resonance energy transfer (FRET) applications, as biosensors in  
 CC prokaryotic and eukaryotic cells, in screening assays, as second  
 CC messenger detectors, in fluorescence activated cell sorting applications,  
 CC in protease cleavage assays or as fluorescent timers. The present  
 CC sequence is a DNA encoding Diacosoma sp. dFp583/dmFp592 (6/90) protein  
 of the invention

XX Sequence 678 BP; 145 A; 217 C; 218 G; 98 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	1.36e-128	Length:	678
Score:	1119.00	Matches:	206
Percent Similarity:	96.89%	Conservative:	12
Best Local Similarity:	91.56%	Mismatches:	7
Query Match:	92.17%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AAD46280 (1-678)

```

OY 1 MetArgSerSerLYaAnValIleYsgLuphMetAArgPheLYaIArgMetGluGly 20
DB 1 ATGAGCTGCAGCAAGAGATGATCAAGAGATTCAATCGGTTCAAGTCCGATGAGGAC 60
OY 21 ThrValaAngLYHAGLuphGluIleGluGluGluGluGluGluGluGluGluGluGlu 40

```

```

Db      61 ACCGTGAACGGCCACGAGTTTCAGATCAAGGCGGAGGGCGGAGCGCCCTTACGAGGCG 120
Qy      41 HiebsanthrVallysleuLysValThrLysGlyGlyProLeuProPheLeuATTPAaplle 60
Db      121 CACTGACAGCGGTGAAGCTCATGTGTACCAAGGCGGCGCCCTCCCTTCGCTTCGATCATC 180
Qy      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAPR11aPro 80
Db      181 CTGAGCCCCCGATTCAGTACGAGCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 240
Qy      81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTyrGluArgValMetAsnPheGlu 100
Db      241 GACTACAAAGAAAGCTCAGCTTCCCGAGGGCTTCAAGTGGAGCGGGGTGATGAATTCTGAG 300
Qy      101 AspglyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyPheIleTyr 120
Db      301 GACGCGCGCGTGTGACCGTGAAGCCAGACGACGCTCAAGAGACGGCTGCTTCATCTAC 360
Qy      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db      361 GAGGTGAAGTTCAATCGCGGTGAATCTTCCCGACGACGGCCCGTGAATGACGGCGGAC 420
Qy      141 MetGlyTyrGlnLysSerThrGluArgLeuTyrProArgAspGlyValLysLysGlyGlu 160
Db      421 CGGGGCTGGAGGCGCAGAGAGCGGCTTACCCCGGAGCGGCGGTCTCAAGGCGGAC 480
Qy      161 LLeHisLysAlaLeuLysLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db      481 ATCCACATGGCCCTCCGCGCTCGAAGGCGGCGGCCACCTACCTCGTGAAGTTCAAGACATC 540
Qy      181 TyrMetLysLysLysProValGlnLeuProGlyTyrTyrTyrValLysSerLysLeuAsp 200
Db      541 TACATGGCCCAAGAACCCCGTGCAGCTTCCCGGCTACCTACCTGAGACGAGACCTGCAC 600
Qy      201 LLeHisSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGlnGlyArg 220
Db      601 ATCCACGAGCCCAAGAGCTACACCATCGTGAAGCAGTACGAGCGGACCGAGGCGCG 660
Qy      221 HisHisLeuPheLeu 225
Db      661 CACCACCTCTTCTC 675

```

RESULT 74

AAA48749

ID AAA48749 standard; cDNA; 678 BP.

AC AAA48749;

XX 19-SEP-2000 (first entry)

XX Humanised Anthozoa novel fluorescent protein dfrp583/dmfp592-2G cDNA.

DE Anthozoa; dfrp583; fluorescent protein; non-bioluminescent organism;

KW fluorescent labeling; hybrid; mutant; mutain; mutain; ss.

XX Diacosome sp; "red".

OS Synthetic.

XX Key Location/Qualifiers

FT 1..678

CD CDS /product= "dfrp583/dmfp592-2G mutant hybrid protein"

XX W0200034326-A1.

XX 15-JUN-2000.

XX 10-DEC-1999; 99WO-US029473.

XX 11-DEC-1998; 98US-00210330.

XX 14-OCT-1999; 99US-00418529.

XX (CLON-) CLONTECH LAB INC.

XX

PI Lukyanov SA, Fradkov AF, Labas VA, Matz MV, Green G, Chen Y;  
PI Ding L;  
XX WPI; 2000-423381/36.  
DR P-PSDB; AAY9839.  
XX

PT Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
XX

PS Claim 6, Page 79; 86pp; English.

XX

CC The present sequence is derived from the nucleotide sequence of humanised  
CC hybrid dfrp583/dmfp592. dfrp583 is a full-length cDNA encoding a novel  
CC fluorescent protein (nfp) from Discosoma sp. "red", a non-bioluminescent  
CC species of the Class Anthozoa. The hybrid sequence was changed to  
CC optimise the codons for expression of the protein in mammalian cells. The  
CC present sequence encodes two substitutions, K15Q and I217S, which were  
CC incorporated by random mutagenesis. Fluorescent proteins can be used in  
CC fluorescent labeling, a useful tool for marking a protein, cell or  
CC organism of interest. Unlike other markers used in protein labeling, such  
CC as beta-galactosidase and luciferase, fluorescent proteins do not require  
CC an exogenous cofactor or substrate. Methods involving fluorescent  
CC proteins are also less laborious and less difficult to control than the  
CC traditional methods of fluorescent labeling, where a protein of interest  
CC is purified and then covalently conjugated to a fluorophore derivative.  
CC Novel fluorescent proteins isolated from species of the Class Anthozoa  
CC can be used as markers for gene expression and protein localization  
CC studies, and in fluorescence resonance energy transfer (FRET) reactions.  
CC They may have improved properties and better suitability for larger  
CC excitations compared to prior art fluorescent proteins such as green  
CC fluorescent protein

CC

XX

SQ Sequence 678 BP; 143 A; 218 C; 218 G; 99 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1..33e-127	Length:	678
Score:	1111.00	Matches:	204
Percent Similarity:	96.89%	Conservative:	14
Best Local Similarity:	90.67%	Mismatches:	7
Query Match:	91.52%	Indels:	0
DB:	3	Gaps:	0

US-10-006-922A-12 (1-225) x AAA48749 (1-678)

```

Qy      1 MetArgSerSerLysAsnValLLeLysGluPheMetArgPheLysValArgMetGlnGly 20
Db      1 ATGAGCTGCAGCAAGACGTGATCAAGAGTTTATGCGGTTCAGAGGCGGATGGAAGGCG 60
Qy      21 ThrValAsnGlyHisGluPheGlnIleGlnGlyGlyGlyProLeuProTyrGlnGly 40
Db      61 ACCGTGAACGGCCACGAGTTTCAGATCAAGGCGGAGGGCGGAGCGCCCTTACGAGGCG 120
Qy      41 HiebsanthrVallysleuLysValThrLysGlyGlyProLeuProPheLeuATTPAaplle 60
Db      121 CACTGACAGCGGTGAAGCTCATGTGTACCAAGGCGGCGCCCTCCCTTCGCTTCGATCATC 180
Qy      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAPR11aPro 80
Db      181 CTGAGCCCCCGATTCAGTACGAGCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 240
Qy      81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTyrGluArgValMetAsnPheGlu 100
Db      241 GACTACAAAGAAAGCTCAGCTTCCCGAGGGCTTCAAGTGGAGCGGGGTGATGAATTCTGAG 300
Qy      101 AspglyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyPheIleTyr 120
Db      301 GACGCGCGCGTGTGACCGTGAAGCCAGACGACGCTCAAGAGACGGCTGCTTCATCTAC 360
Qy      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db      361 GAGGTGAAGTTCAATCGCGGTGAATCTTCCCGACGACGGCCCGTGAATGACGGCGGAC 420
Qy      141 MetGlyTyrGlnLysSerThrGluArgLeuTyrProArgAspGlyValLysLysGlyGlu 160

```

```

Db      421 CGGGGCTGGAGGCGCACGACGGCTCTAACCCCGGAGCGGCTCAAGGCGAC 480
Qy      161 ILeHsYsAlaLeuYsLeuYsAAspGlyGlyYHsYrLeuValGluPheSerIle 180
Db      481 ATCCACATGGCCCTCCGAGCTCGAGGCGGCGCACTACCTCGTGGAGTTCAAGCATC 540
Qy      181 TyMeAlaYsYsProValGluLeuProGlyTyrTyrTyrValAAspSerYsLeuAAsp 200
Db      541 TACATGGCCAAAGAGCCCGTGCAGCTCCCGGCTACTACTACGTGACAGCAAGCTGCAC 600
Qy      201 ILeHsSerHsAAspGlyAAspTyrThrIleValGluGlnTyrGluAAspThnGluYArg 220
Db      601 ATCCACAGCCCAACAGAGACTACCATCGTGGAGCAAGTACGAGCGGTCCGAGGCGCG 660
Qy      221 HsHsLeuPheLeu 225
Db      661 CACCACTCTTCTCTC 675

RESULT 75
AAA48750
ID AAA48750 standard; cDNA; 678 BP.
AC AAA48750;
XX
XX 19-SEP-2000 (first entry)
DE Humanised Anchozoa novel fluorescent protein drFP583/dmFP592-Q3 cDNA.
XX
XX Anchozoa; drFP583; fluorescent protein; non-bioluminescent organism;
XX fluorescent labeling; hybrid; mutant; muteln; ss.
XX
XX Diacosoma sp; "red".
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 1..678
XX /product= "drFP583/dmFP592-Q3 mutant hybrid protein"
XX
XX WO200034326-A1.
XX
XX 15-JUN-2000.
XX
XX 10-DEC-1999; 99WO-US029473.
XX
XX 11-DEC-1998; 98US-00210330.
XX 14-OCT-1999; 99US-00418529.
XX
XX (CLON-) CLONTECH LAB INC.
XX
XX Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;
XX Ding L;
XX
XX WPI; 2000-423381/36.
XX P-PSDB; AAY99840.
XX
XX Novel fluorescent protein from non-bioluminescent Diacosoma sp. red,
XX useful for fluorescent labeling and as markers.
XX
XX Claim 6, Page 80-81; 86pp; English.
XX
XX The present sequence is derived from the nucleotide sequence of humanised
XX hybrid drFP583/dmFP592. drFP583 is a full-length cDNA encoding a novel
XX fluorescent protein (nfp) from Diacosoma sp. "red", a non-bioluminescent
XX species of the Class Anchozoa. The hybrid sequence was changed to
XX optimise the codons for expression of the protein in mammalian cells. The
XX present sequence encodes three substitutions, K15Q, K83M and T217S, which
XX were incorporated by random mutagenesis. Fluorescent proteins can be used
XX in fluorescent labeling, a useful tool for marking a protein, cell or
XX organism of interest. Unlike other markers used in protein labeling, such
XX as beta-galactosidase and luciferase, fluorescent proteins do not require
XX an exogenous cofactor or substrate. Methods involving fluorescent
XX proteins are also less laborious and less difficult to control than the

```

```

CC traditional methods of fluorescent labeling, where a protein of interest
CC is purified and then covalently conjugated to a fluorophore derivative.
CC Novel fluorescent proteins isolated from species of the Class Anchozoa
CC can be used as markers for gene expression and protein localization
CC studies, and in fluorescence resonance energy transfer (FRET) reactions.
CC They may have improved properties and better suitability for larger
CC excitations compared to prior art fluorescent proteins such as green
CC fluorescent protein
XX
XX SQ Sequence 678 BP; 142 A; 218 C; 218 G; 100 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 7,36e-127 Length: 678
XX Score: 1105.00 Matches: 203
XX Percent Similarity: 96.44% Conservative: 14
XX Best Local Similarity: 90.22% Mismatches: 8
XX Query Match: 91.02% Indels: 0
XX DB: 3 Gaps: 0
XX
XX US-10-006-922a-12 (1-225) x AAA48750 (1-678)
Qy      1 MeArGSeSerYsAenValIleYsGluPheMeArGpHeYsValArgMeTcGluGly 20
Db      1 ATGAGCTGAGCAAGAGAGAGTGTACAGAGATTCAATCGGTTCCAGTGCAGATGAGAGGC 60
Qy      21 ThrValAsnGlyYHsGluPheGlnIleGluGlyGluYArgProTyrGluGly 40
Db      61 ACCGTGAACGGCCACGAGTTGAGATCAAGGCGCAGAGGCGGCGCCCTTACAGAGGCG 120
Qy      41 HsAsnThrValYsLeuYsValThrYsGlyGlyProLeuProPheAlaTyrAspIle 60
Db      121 CACTGACGCGGAAGCTCAGTACGTAACCAAGGCGGCGCCCTCCCTTCGCTTGACATC 180
Qy      61 LeuSerProGlnPheGlnTyrGlySerYsValTyrValYsHsIProAlaAspIlePro 80
Db      181 CTGAGCCGCCAGTTCAGTACGCGCAGCAAGGTGTACGTAAGCAACCCGCGACATCCCC 240
Qy      81 AspTyrYsYsLeuSerPheProGluGlyPheYsTrpGluValMetAsnPhnGlu 100
Db      241 GACTACATGAAGCTCAGCTTCCCGAGGCGCTTCAAGTGGGCGCGGTATGACTTCAG 300
Qy      101 AspGlyGlyValValThrValThrGluAspSerSerLeuGlnAAspGlyCysPheIleTyr 120
Db      301 GACGCGCGCGTGTGACCGTGAAGCAGAGACGAGCTTCMAAGACGCGCTTCATCTAC 360
Qy      121 LySValYsPheIleGlyValAsnPheProSerAspGlyProValMetGlnYsYsThr 140
Db      361 GAGGTGAAGTTCACTCGGTGAATCTCCCGACGACGCGCCCGTATGACAGCGGAGACC 420
Qy      141 MetGlyTrpGluAAspThrGluArgLeuTyrProArgAspGlyValLeuYsGlyGlu 160
Db      421 CGGGGCTGGAGGCGACGACGGCTCTAACCCCGGAGCGGCTCAAGGCGAC 480
Qy      161 ILeHsYsAlaLeuYsLeuYsAAspGlyGlyYHsYrLeuValGluPheSerIle 180
Db      481 ATCCACATGGCCCTCCGAGCTCGAGGCGGCGCACTACCTCGTGGAGTTCAAGCATC 540
Qy      181 TyMeAlaYsYsProValGluLeuProGlyTyrTyrTyrValAAspSerYsLeuAAsp 200
Db      541 TACATGGCCAAAGAGCCCGTGCAGCTCCCGGCTACTACTACGTGACAGCAAGCTGCAC 600
Qy      201 ILeHsSerHsAAspGlyAAspTyrThrIleValGluGlnTyrGluAAspThnGluYArg 220
Db      601 ATCCACAGCCCAACAGAGACTACCATCGTGGAGCAAGTACGAGCGGTCCGAGGCGCG 660
Qy      221 HsHsLeuPheLeu 225
Db      661 CACCACTCTTCTCTC 675

RESULT 76
AAA28958
ID AAA28958 standard; cDNA; 876 BP.
XX

```

AA28958;  
12-SEP-2000 (first entry)  
Discosoma sp. "magenta" dmfp592 cDNA encoding a fluorescent protein.  
dmfp592; fluorescent protein; non-bioluminescent; label; marker;  
fluorescence resonance energy transfer reaction; FRET; ss.  
Discosoma sp.  
Key Location/Qualifiers  
CDS 45..737  
/\*tag= a  
/product= "fluorescent\_protein"  
MO200034324-A1.  
15-JUN-2000.  
10-DEC-1999; 99MO-US029412.  
11-DEC-1998; 98US-00210330.  
19-NOV-1999; 99US-00444341.  
(CLON-) CLONTECH LAB INC.  
Lukyanov SA, Fradkov AF, Labas YA, Matz MV;  
WPI; 2000-423379/36.  
P-PSDB; AAY92894.  
Novel fluorescent protein from non-bioluminescent Discosoma species,  
useful for fluorescent labeling and as markers.  
Claim 7; Page 54-55; 60pp; English.  
This cDNA encodes a fluorescent protein (dmfp592) from Discosoma sp.  
"magenta", a non-bioluminescent species of the genera Anthozoa. The  
fluorescent colour is observed under usual white light or ultraviolet  
light. It has an absorbance maximum at 573 nm and an emission maximum at  
593 nm. The polynucleotides of the invention can be used to produce the  
proteins recombinantly, and as a source of primers and probes for  
identifying related proteins. The fluorescent proteins have applications  
in fluorescent labeling, as fluorescent markers for gene expression and  
protein localization studies, and in fluorescence resonance energy  
transfer (FRET) reactions. They may have improved properties and better  
suitability for larger excitations, compared to prior art fluorescent  
proteins such as green fluorescent protein from *Aequorea victoria*, which  
has an excitation maximum at 395 nm, a second excitation peak at 475 nm  
and an emission maximum at 510 nm  
Sequence 876 BP; 256 A; 165 C; 232 G; 223 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2.75e-124 Length: 876  
Score: 1085.50 Matches: 199  
Percent Similarity: 95.58% Conservative: 17  
Best Local Similarity: 88.05% Mismatches: 9  
Query Match: 89.42% Indels: 1  
DB: 3 Gaps: 1  
US-10-006-922a-12 (1-225) x AA28958 (1-876)  
QY 1 MetArgSerSerIyAsnValIleLyseGluPheMetArgPheLyValArgMetGluGly 20  
DB 45 ATGAGTGTTCCTCAAGAAATGATCATCAAGAGTTCTCAAGTTCCGATGAAAGGA 104  
QY 21 ThrValangIyHsGluPheGluIleGluGlyGluGlyGluGlyGluGlyGluGly 40  
DB 105 ACGGTCAATGGGCGACGAGTTTGAATAAAGCCAGAGGTGAAGGAGCCCTTACGAAGGT 164  
QY 41 HisAnthrValLybLeuLyValThrLyseGlyGlyProLeuProPheAlaTrpAspIle 60

DB 165 CACTGTCCTCGTAAGCTTANGTGAACCAAGGTGACCTTTGCATTTGTTGATATT 224  
QY 61 LeuSerProGlnPheGlnTrpGlySerLyValYrValLybHisProAlaAspIlePro 80  
DB 225 TTGTCAACCAATTTCAGTATGGAACCAAGTATATGTCAACACCTGCGCATACCA 284  
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLyTrpGluLybValMetAsnPheGlu 100  
DB 285 GACTATTAAGAGCTGATTCCTGAGGAGATTAAATGGGAAGGTCATGAACCTTGAA 344  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyPheIleYr 120  
DB 345 GACGGTGGCGTGTACTGATATCCAAATTCAGTTGAAGACGCGCTGTTCACTTAC 404  
QY 121 LysValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybTrp 140  
DB 405 GAGGTCAAGTTCATTCGGGGTGAACCTTCCTGATGAGACCTGTTATGCAAGAGAGCA 464  
QY 141 MetGlyTrpGluLybSerThrGluArgLeuYrProArgAspGlyValLeuLybGlyLu 160  
DB 465 CGGGGCTGGAGACCGAGCTTGAGCCGTTGTATCTTGATGGGTGCTGMAAGAGAC 524  
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisYrLeuValGluPheLybSerIle 180  
DB 525 ATCCATATGCTCTGAGGCTGGAAGAGGCGGCATTTACCTGTTGAATTCAAAAGTATT 584  
QY 181 TyrMetAlaLybLybPro--ValGlnLeuProGlyYrTyrTyrYrValAspSerLybLeu 199  
DB 585 TACATGTATTAAGAAAGGCTTCAGTGCAAGTTCAGGCTACTATTATATGTGCTCAACTG 644  
QY 200 AspIleThrSerHisAngrLybAspYrThrIleValGluGlnTrpGluArgThrGluGly 219  
DB 645 GATATGACGAGCGACCAAGAAATTCACAGTCTTGAGCAGTATGAAAAAACCCAGGGA 704  
QY 220 ArgHisHisLeuPheLeu 225  
DB 705 CGCCACCATCCGTTCAATT 722  
RESULT 77  
AAD03618  
ID AAD03618 standard; cDNA; 876 BP.  
XX  
AC AAD03618;  
XX  
XX 11-SEP-2003 (revised)  
DT 19-JUN-2001 (first entry)  
XX  
DE Discosoma sp. magenta chromo/fluorescent protein, dmfp592 (NFP-9) cDNA.  
XX  
KW Anthozoa; Chromoprotein; fluorescent protein; dmfp592; NFP-9; sunscreen;  
KW analyte detection assay; selectable marker; recombinant DNA application;  
KW biosensor; pH indicator; invivo marker; selective filter; ss.  
XX  
OS Discosoma sp; magenta.  
XX  
FH Key Location/Qualifiers  
FT CDS 45..737  
FT /\*tag= a  
FT /product= "Chromo/fluorescent protein, dmfp592 (NFP-9) "  
XX  
PN MO200127150-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 13-OCT-2000; 2000MO-US028477.  
XX  
PR 14-OCT-1999; 99US-00418529.  
PR 15-OCT-1999; 99US-00418917.  
PR 15-OCT-1999; 99US-00418922.  
PR 19-NOV-1999; 99US-00444338.  
PR 19-NOV-1999; 99US-00444341.  
PR 09-DEC-1999; 99US-00457556.

PR 09-DEC-1999; 99US-00457898.  
 PR 09-DEC-1999; 99US-00458144.  
 PR 09-DEC-1999; 99US-00458477.  
 PR 10-DEC-1999; 99MO-US029405.  
 PR 14-JUN-2000; 2000US-0211607P.  
 PR 14-JUN-2000; 2000US-0211609P.  
 PR 14-JUN-2000; 2000US-0211626P.  
 PR 14-JUN-2000; 2000US-0211627P.  
 PR 14-JUN-2000; 2000US-0211687P.  
 PR 14-JUN-2000; 2000US-0211766P.  
 PR 14-JUN-2000; 2000US-0211888P.  
 PR 14-JUN-2000; 2000US-0211888P.  
 PR 14-JUN-2000; 2000US-0212070P.  
 XX  
 PA (CLON-) CLONTECH LAB INC.  
 XX  
 P1 Lukyanov SA, Fradkov AF, Labas YA, Matcz MV, Terexikh A;  
 XX  
 XX WPI; 2001-266409/27.  
 DR P-PSDB; AAE00378.  
 DR  
 PT An Anchozoa chromo- or fluorescent protein (PI) present in an environment  
 PT other than its natural environment, useful as a label in analyte  
 PT detection assays and as a selectable marker in recombinant DNA  
 PT applications.  
 PT  
 XX  
 PS Claim 13; Fig 9; 69pp; English.  
 XX  
 CC The present sequence is a Discosoma sp. magenta chromo/fluorescent  
 CC protein, dmpP592 (NFP-9) cDNA. NFP-9 is present in an environment other  
 CC than its natural environment and has an absorbance maximum ranging from  
 CC 375nm to 775nm and more usually from 560nm to 590nm and emission maximum  
 CC ranging from 395nm to 795nm and more usually from 580 to 610nm. The  
 CC chromoproteins or fluorescent proteins are useful as labels in analyte  
 CC detection assays, as selectable markers in recombinant DNA applications,  
 CC as bioemore in prokaryotic and eukaryotic cells e.g. as pH indicator  
 CC and as in vivo markers in animals. They are also useful in sunscreens and  
 CC as selective filters. (Updated on 11-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 876 BP; 256 A; 165 C; 232 G; 223 T; 0 U; 0 Other:  
 Alignment Scores:  
 Pred. No.: 2,75e-124 Length: 876  
 Score: 1085.50 Matches: 199  
 Percent Similarity: 95.58% Conservative: 17  
 Best Local Similarity: 88.05% Mismatches: 9  
 Query Match: 89.42% Indels: 1  
 DB: 4 Gaps: 1  
 US-10-006-922a-12 (1-225) x AAD03618 (1-876)  
 QY 1 MetAgsSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
 DB 45 ATGAGTTGTTCCAGAGATGTGATCAAGAGTTCAAGAGTTCAAGAGTTCAAGAGATG 104  
 QY 21 ThrValaenGlyHisGluPheGluIleGluGlyGluGlyGlyGlyGlyGlyGlyGly 40  
 DB 105 ACGGTCAATGGGACGAGTTTGAATTAAGCGGAAGGTGAAGGAGGCTTACGAAGGT 164  
 QY 41 HisaenThrValaenLysLysValThrLysGlyGlyGlyGlyGlyGlyGlyGlyGly 60  
 DB 165 CACTGTTCCGTAAAGCTTATGTAACCAAGGAGGCTTTGCCATTTCTTTGATATT 224  
 QY 61 LeuSerProGlnPheGlnIleGlySerLysValIleValIleHisProLysPhePro 80  
 DB 225 TTGTCAACCAATTCAGATGAGAACAGGATATGTAACACCCCTCCGACATACCA 284  
 QY 81 AAGTATLysLysLeuSerPheProGluGlyPheLysTrpGluLysValMetAsnPheGlu 100  
 DB 285 GACTATTAAGAGTGTCAATTTCTGTGAGGATTTAAATGGGAAGGCTCATGAACCTTGA 344  
 QY 101 AAGGAGLysLysValaenThrValaenPheSerLeuGlnAspGlyCysPheIleTyr 120

DB 345 GACGTTGCGGTGTACTGATTCAGATTCAGATTGAAAGACGCGCTGTTCACTAC 404  
 QY 121 LysValaenPheIleGlyValaenPheProSerAspGlyProValMetGlnLysLysThr 140  
 DB 405 GAGGTCAAGTTCAATGGGGTGAACCTTCTTCGATGAGACCTGTATGACAGAGAGACA 464  
 QY 141 MetGlyTrpGluLysSerThrLysLysLysLysLysLysLysLysLysLysLysLys 160  
 DB 465 CGGGGCTGGGAAGCAGCTGAGCCGTTGTATCTCTGTATGGGGCTGTAAGAGAGAC 524  
 QY 161 ILeHisLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 180  
 DB 525 ATCCATATGCTCTGAGAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584  
 QY 181 TyrMetAlaLysLysPro---ValGlnLeuProGlyTyrTyrTyrValaenSerLysLeu 199  
 DB 585 TACATGGTAAAGAACCTTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG 644  
 QY 200 AspIleThrSerHisAsnGluAspTyrThrIleValGluGlnIleGluLysGluGly 219  
 DB 645 GATATGACGAGCCACCAAGAAATTAACAGCTGTGAGAGATTAAGAAACCAGGGA 704  
 QY 220 ArgHisLysLeuPheLeu 225  
 DB 705 GCGCACCATCGTTCATT 722  
 RESULT 78  
 ACC44649  
 ID ACC44649 standard; DNA; 876 BP.  
 XX  
 AC ACC44649;  
 XX  
 DT 29-MAY-2003 (first entry)  
 DE  
 XX Red fluorescent protein (DeRed) encoding DNA SEQ ID NO:39.  
 XX  
 KM Chromosome-based platform; artificial chromosome; eukaryotic chromosome;  
 KM att site; integrase; recombinase; Aces; gene therapy; transgenic animal;  
 KM platform artificial chromosome expression system; gene; ds.  
 XX  
 OS Discosoma sp.  
 OS  
 OS Synthetic.  
 XX  
 PN WO200297059-A2.  
 XX  
 PD 05-DEC-2002.  
 XX  
 PP 30-MAY-2002; 2002MO-US017452.  
 XX  
 PR 30-MAY-2001; 2001US-0294758P.  
 PR 21-MAR-2002; 2002US-0366891P.  
 PA (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.  
 PI Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;  
 PI Stewart S, Shellard J;  
 DR WPI; 2003-140461/13.  
 DR P-PSDB; ABP96651.  
 XX  
 PT Novel eukaryotic chromosome comprising one or many att sites which  
 PT permits site-directed integration in the presence of lambda-integrase,  
 PT useful for site-specific recombination-directed integration of DNA of  
 PT interest.  
 PS Example 1; Page 191-192; 272pp; English.  
 XX  
 CC The present invention describes a eukaryotic chromosome (I) comprising  
 CC one or several att sites, where an att site is heterologous to the  
 CC chromosome, and permits site-directed integration in the presence of  
 CC lambda-integrase. Also described: (1) a platform artificial chromosome  
 CC expression system (Aces) (II) comprising several sites that participate  
 CC in recombinase catalysed recombination; and (2) a method (M1) for

CC introducing a heterologous nucleic acid into a platform artificial  
 CC chromosome. (I) can be used in gene therapy. (M1) is useful for  
 CC introducing a heterologous nucleic acid molecule into a platform  
 CC artificial chromosome, preferably an Aces. (II) is useful for producing a  
 CC transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or  
 CC mammal) by introducing (II) by cell fusion, lipid-mediated transfection  
 CC by a carrier system, microinjection, microcell fusion, electroporation,  
 CC microparticle bombardment or direct DNA transfer into an embryonic  
 CC cell, preferably a stem cell or an embryo. (II) comprises a heterologous  
 CC nucleic acid that encodes a therapeutic product which is useful for  
 CC making a library of Aces comprising random portions of a genome. ACC44612  
 CC to ACC44732 and ABP96650 to ABP96657 represent sequences used in the  
 CC exemplification of the present invention

XX  
 SQ Sequence 876 BP, 256 A, 165 C, 232 G, 223 T, 0 U, 0 Other;

Alignment Scores:  
 Pred. No.: 2,75e-124 Length: 876  
 Score: 1085.50 Matches: 199  
 Percent Similarity: 95.58% Conservative: 17  
 Best Local Similarity: 88.05% Mismatches: 9  
 Query Match: 89.42% Indels: 1  
 Gaps: 1

US-10-006-922A-12 (1-225) x ACC44649 (1-876)

QY 1 MetArgSerSerLybAnValIleLygGluPheMetArgPheLybValArgMetGluGly 20  
 DB 45 ATGAGTGTTCACAGAAATGTCATCAGAGATTCAGAGTTCAGAGTTCAGATGGAAGA 104  
 QY 21 ThrValAnGlyYH1sgLupPheGluIleGluLygGluGluYAcProTyGluGly 40  
 DB 105 ACGGTCAATGGGCGAGAGTTTGAAATTAAGGCAAGGTGAAGGAGGCTTACGAAGGT 164  
 QY 41 HisAnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60  
 DB 165 CACGTTCGCGAAGCTTATGATGTAACCAAGGCTGACCTTGGCATTTGCTTTGATATT 224  
 QY 61 LeuSerProGlnPheGlnTyGlySerLybValTyValLybSH1sProAlaAspIlePro 80  
 DB 225 TTGTCCACACAAATTCAGTATGAGAAAGCAAGGTATATGCAACACCTGCGACATACCA 284  
 QY 81 AspTyTrLybLybLeuSerPheProGluGlyPheLybTrpGluAcValMetAsnPheGlu 100  
 DB 285 GACATATAAAGCTGTCAATTCCTGAGGGAATTTAAATGGGAAAGGCTCATGAACTTTGAA 344  
 QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTy 120  
 DB 345 GAGGTGGCGGTGATCGTATCCCAAGATTCAGTTGAAAGAAGGCTGTTTCATCTAC 404  
 QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybThr 140  
 DB 405 GAGGTCAAGTTCATTTGGGGTGAACCTTCCTCATGGAAGCTGTTATGACAGAGAGACA 464  
 QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyProArgAspGlyValIleLeuGlyGlu 160  
 DB 465 CGGGAGCTGGAGACCAAGCTCTGAGCGTTTGATCTCGTATGGGGTGTCTAAAGAGAC 524  
 QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyLeuValGluPheLybSerIle 180  
 DB 525 ATCATATGGCTCGAGGCTGGAAGGAGGCGCATTAACCTGTTGAATTCAAAAGATTT 584  
 QY 181 TyrMetAlaLybLybPro---ValGlnLeuProGlyTyTrTyTrValAspSerLybLeu 199  
 DB 585 TACATGTGAAAGAGCTTCAGTCAGTTCACAGGCTACTTATATGTTAGCTCCAAACTG 644  
 QY 200 AspIleThrSerSH1sAnGluAspTyThrIleValGluGlnTyArgIuArgThrGluGly 219  
 DB 645 GATATGAGCAAGCAACAGAAATTAACAGTGTGAGAGTATGAATAAACCAAGGGA 704  
 QY 220 ArgHisHisLeuPheLeu 225  
 DB 705 CGCCACCATCGCTTCATT 722

RESULT 79  
 ADC24132  
 ID ADC24132 standard; DNA; 678 BP.  
 XX  
 AC ADC24132;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Discosoma red fluorescent protein variant mRFP1 DNA.  
 XX  
 KM Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
 KM fluorescence protein variant; transcription induction detection;  
 KM fluorescence energy resonance transfer; FRET; protein kinase;  
 KM protein phosphatase; ion indicator; mutant; mRFP1; ds.  
 XX  
 OS Synthetic.  
 OS Discosoma.  
 XX  
 PN US2003059835-A1.  
 XX  
 PD 27-MAR-2003.  
 XX  
 PF 10-APR-2002; 2002US-00121258.  
 XX  
 PR 26-FEB-2001; 2001US-00794308.  
 PR 24-MAY-2001; 2001US-00866538.  
 PA (TSIE/) TSIE R Y.  
 PA (CAMP/) CAMPBELL R E.  
 PI Tsien RY, Campbell RE;  
 XX  
 DR WPI, 2003-743764/70.  
 DR P-PSDB; ADC24133.  
 XX  
 PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
 PT variant having a reduced propensity to oligomerize, useful for detecting  
 PT transcriptional activity.  
 XX  
 PS Disclosure; SEQ ID NO 7; 678bp; English.  
 XX  
 CC The invention describes a polynucleotide sequence (I) encoding a  
 CC Discosoma red fluorescent protein (DsRed) variant having a reduced  
 CC propensity to oligomerize, comprising amino acid substitutions at the AB  
 CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
 CC amino acids, given in the specification, where the substitutions result  
 CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
 CC useful for detecting transcriptional activity by providing a host cells  
 CC containing a vector which comprises (i) operatively linked to an  
 CC expression control sequence, and an unit to assay the variant fluorescent  
 CC protein fluorescence, and assaying fluorescence of the variant  
 CC fluorescent protein produced by (VII), where variant fluorescent protein  
 CC fluorescence is indicative of transcriptional activity. A polynucleotide  
 CC encoding a fusion protein is useful for the analysis of in vivo  
 CC localisation or trafficking of a polypeptide of interest. A polypeptide  
 CC marker is useful as markers to identify the location and amount of a  
 CC target protein produced, where the target protein is fused to the marker,  
 CC as a complement to or alternative for the green fluorescent protein or  
 CC its spectral variant, for detecting induction of transcription, in  
 CC applications involving fluorescence energy resonance transfer (FRET),  
 CC which detects events as the function of the movement of fluorescent  
 CC donors and acceptors towards or away from each other, for making  
 CC fluorescent sensors for protein kinase and phosphatase activities or  
 CC indicators for ions and molecules such as Ca<sup>2+</sup>, Zn<sup>2+</sup>, for identifying the  
 CC presence of a molecule in a sample, for identifying a specific  
 CC interaction of a first and second molecule, for determining whether a  
 CC sample contains an enzyme or for determining the pH of the sample. (I) is  
 CC useful for identifying a region or condition that regulates the activity  
 CC of an expression control sequence. This sequence encodes Discosoma red  
 CC fluorescent protein variant mRFP1.  
 XX  
 SQ Sequence 678 BP, 149 A, 228 C, 208 G, 93 T, 0 U, 0 Other;



## Alignment Scores:

Pred. No.: 1,88e-116 Length: 678  
 Score: 1021.00 Matches: 192  
 Percent Similarity: 89.14% Conservative: 5  
 Best Local Similarity: 86.88% Mismatches: 24  
 Query Match: 84.10% Indels: 0  
 Gaps: 0

US-10-006-922a-12 (1-225) x ADL46211 (1-678)

```

QY 1 MetArgSerSerlybAenValIlelysgluPhenetaRphelybValArgMetGluGly 20
DB 1 ATGGCCCTCCTCCGAGAGCTGATCAAGAGATTCTGCGCTTCAAGGTGCGCATGAAAGGC 60
QY 21 ThrValaAnglybHsgluPhenGluIleGluGlyGluGlyGluGlyArgProTyGlyGly 40
DB 61 TCCGTGAACGGCCACGAGTTCAAGTTCAGAGGCGAGGGCCGAGGGCCGCTTACGAGGGC 120
QY 41 HisAenThrVallybLeuLybValThrlysglyGlyProLeuProPheAlaTrpAspIle 60
DB 121 ACCGAGACCGCCAGCAAGCTGAAGGTGACCAAGGCGGCCCCCTGCTTGCCTGGGACATC 180
QY 61 LeuSerProGluPheGluTrpGlySerlybValTyVallybHisProAlaAspIlePro 80
DB 181 CTGTCCCTCAGTTCAGTACGCGCTCCAGAGCTTACGTAAGCAACCCGCGACATCCCC 240
QY 81 AspTyTrpLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPhenGlu 100
DB 241 GACTACTTGAAAGCTGTCTCTCCGAGGGCTTCAAGTTCAGAGGCGGCTGATGACTTCAG 300
QY 101 AspGlyGlyVallybValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTy 120
DB 301 GACGGCGCGCGTGTGACCGGTGACCCAGACCTCTCCCTGCAAGCAAGGCGAGTTCATCAC 360
QY 121 LybVallybPheIleGlyValaAsnPhenProSerAspGlyProValMetGlnLybLybThr 140
DB 361 AAGGTGAAGCTGCGCGGCAACCTTCCCTCCGAGGCGCCCGTAAATGCAGAAAGAGCC 420
QY 141 MetGlyTrpGluAspThrGluArgLeuTyTrpProArgAspGlyValIleuLybGlyGlu 160
DB 421 ATGGGCTGGAGGCTTCCACGAGGGAGTACCCCGAGAGCGGCGCTTGAAGGCGAG 480
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyTrpLeuValGluPheLybSerIle 180
DB 481 ATCAAGATGAGGCTGAAGCTGAAGGACGGCGGCACTACGACGCGAGATCAAGACCAAC 540
QY 181 TyrMetAlaLybLybPheProValGlnLeuProGlyTyTrpTyTrpValAspSerlybLeuAsp 200
DB 541 TACATGGCCCAAGAGCCCGGTGACGCTCCGCGCTTCAAGACCGGACATCAAGCTGAGC 600
QY 201 IleThrSerHisAenGluAspTyTrpThrIleValGluGlnTyTrpGluArgThrGluGlyArg 220
DB 601 ATCACTCTCCACAGAGAGATCACATCGTGAACAGTACGAGCGGCGGCGGAGGGCCGC 660
QY 221 His 221
DB 661 CAC 663

```

RESULT 80  
ADL46211  
ID ADL46211 standard; DNA; 678 BP.

AC ADL46211;  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Discosoma red fluorescent protein (DsRed) variant mRFP1 coding sequence.  
 XX red fluorescent protein; DsRed; fluorescence; red wavelength;  
 KM oligomerization; tetramerization; immunosassay; hybridization assay; gene;  
 de.  
 XX

## OS Discosoma sp.

Key Location/Qualifiers  
 CDS 1..678  
 FT /\*tag= a  
 /product= "DsRed variant mRFP1 protein"

W02003086446-A1.

23-OCT-2003.

09-APR-2003; 2003W0-US010879.

10-APR-2002; 2002US-00121258.

29-JUL-2002; 2002US-00209208.

(REGC ) UNIV CALIFORNIA.

Tsien RY, Campbell RE, Baird GS;

WPI; 2003-845265/78.

P-PSDB; ADL46210.

New monomeric and dimeric Anthozoan fluorescent protein variants with reduced propensity to oligomerize, and encoding polynucleotides, useful in molecular biology, e.g. in immunoassays or in tracking protein movement in cells.

Disclosure; SEQ ID NO 9; 16pp; English.

The invention relates to a polynucleotide sequence encoding a Discosoma red fluorescent protein (DsRed) variant having a reduced propensity to oligomerize. The protein variant comprises one or more amino acid substitutions at the AB and/or AC interface(s) of the wild-type DsRed sequence, where the substitutions result in reduced propensity of the DsRed variant to form tetramers and where the variant displays detectable fluorescence of at least one red wavelength. The composition and methods are useful in producing red fluorescent proteins having reduced propensity for oligomerization, especially tetramerization. The protein may be used in molecular biology and in other scientific applications, such as in immunoassays or hybridization assays, or in tracking the movement of proteins in cells. This sequence corresponds to the DsRed variant mRFP1 coding sequence.

Sequence 678 BP; 149 A; 228 C; 208 G; 93 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1,88e-116 Length: 678  
 Score: 1021.00 Matches: 192  
 Percent Similarity: 89.14% Conservative: 5  
 Best Local Similarity: 86.88% Mismatches: 24  
 Query Match: 84.10% Indels: 0  
 Gaps: 0

US-10-006-922a-12 (1-225) x ADL46211 (1-678)

```

QY 1 MetArgSerSerlybAenValIlelysgluPhenetaRphelybValArgMetGluGly 20
DB 1 ATGGCCCTCCTCCGAGAGCTGATCAAGAGATTCTGCGCTTCAAGGTGCGCATGAAAGGC 60
QY 21 ThrValaAnglybHsgluPhenGluIleGluGlyGluGlyGluGlyArgProTyGlyGly 40
DB 61 TCCGTGAACGGCCACGAGTTCAAGTTCAGAGGCGAGGGCCGAGGGCCGCTTACGAGGGC 120
QY 41 HisAenThrVallybLeuLybValThrlysglyGlyProLeuProPheAlaTrpAspIle 60
DB 121 ACCGAGACCGCCAGCAAGCTGAAGGTGACCAAGGCGGCCCCCTGCTTGCCTGGGACATC 180
QY 61 LeuSerProGluPheGluTrpGlySerlybValTyVallybHisProAlaAspIlePro 80
DB 181 CTGTCCCTCAGTTCAGTACGCGCTCCAGAGCTTACGTAAGCAACCCGCGACATCCCC 240
QY 81 AspTyTrpLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPhenGlu 100

```

```
DB 241 GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGATGAACTTGAG 300
QY 101 AaPGLyGLyValValThrValThrGlnAaPSeSerSerleuGlnAaPGLyCySpheileTyr 120
DB 301 GAGGCGCGCGGTGGACCGGTGACCCAGAGACTCCCTCGAGGAGCGCGGATTCATCTAC 360
QY 121 LysValLysPheileGlyValAaPheProSerSerAaPGLyProValMetGlnLysLysThr 140
DB 361 AAGGTGAAGCTGCGCGCACCACTTCCCTCCAGCGCGCCGTAATGCAAGAAAGACC 420
QY 141 MetGLYTrpGlnAaSerThrGlnLysLeuTyrProArgAaPGLyValLleLysGLyGlu 160
DB 421 ATGGGCTGGAGGCTCTCACCGAGCGATGATACCCCGAGAGCGCGCCCTGAAGGGCGAG 480
QY 161 lIeHlalyaLaleuLysLeuLysAaPGLyGLYhIstYrLeuValGluPheLysSerile 180
DB 481 ATCAAGATGAGGCTGAAGCTGAAGACGCGCGGCACTACGACGCGGAGTCAAGACACC 540
QY 181 TyrMetAlalybLysPProValGlnLeuProGLYTrYTrYTrYValAaPSeLysLysAaP 200
DB 541 TACATGCGCAAGAGCGCGGTGACCTGCGCGCTTCAAGACCATCAAGCTGAGC 600
QY 201 lIeThSerHlSaenGluAaPGLYTrThrileValGluGlnTyrGluAaTgThrGluGLYArg 220
DB 601 ATCACTCTCCACAAAGAGACTACACCATCTGGAACATGACGCGCGCGAGGGCGCG 660
QY 221 HIs 221
DB 661 CAC 663

RESULT 81
AaA54911
ID AaA54911 standard; DNA; 2241 BP.
XX AaA54911;
DT 25-AUG-2005 (first entry)
DE Fluorescence-related EGFP-caaspase-3 DEVD linker-mRFP(x2) fusion DNA.
XX fluorescence; protein localization; cellular transport; organelle;
XX mitochondria; de; gene; enhanced green fluorescent protein; EGFP;
XX caaspase-3.
XX Unidentified.
XX OS Chimeric.
XX PN WO2005054464-A1.
XX PD 16-JUN-2005.
XX PF 03-DEC-2004; 2004WO-JP018437.
XX PR 03-DEC-2003; 2003JP-00404472.
XX PR 27-JAN-2004; 2004JP-00018344.
XX PA (RIKE ) RIKEN KK.
XX PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.
XX PI Miyawaki A, Kogure T, Hama H, Kinjo M, Saito K, Karasawa S;
XX Arai T;
XX WPI; 2005-444979/45.
XX DR P-PSDB; AaA54910.
PT Novel fluorescent protein existing as monomer and isolated from Fungia
PT and Montipora species, useful for labeling and analyzing location and
PT movement of intracellular organelles.
XX Example 10; SEQ ID NO 50; 218bp; Japanese.
XX The invention relates to a novel fluorescent protein existing as a
```

```
CC monomer and comprising any one of SEQ ID No. 1-29 (odd SEQ ID numbers)
CC given in the specification, or any one of the sequences chosen from SEQ
CC ID No. 1-29 (odd SEQ ID numbers), with one or more amino acid
CC substitutions, deletions and/or additions, and having fluorescent
CC characteristics. The fusion protein of the invention may be useful for
CC characterizing the location or movement of a protein in a cell in vivo and for
CC labeling and analyzing the location or movement of intracellular
CC organelles, such as mitochondria. The current sequence is that of the
CC fluorescence-related enhanced green fluorescent protein (EGFP)-caaspase-3
CC DEVD linker-mRFP(x2) fusion DNA of the invention.
XX
SQ Sequence 2241 BP; 503 A; 748 C; 670 G; 320 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,01e-115 Length: 2241
Score: 1021.00 Matches: 192
Percent Similarity: 89.14% Conservative: 5
Best Local Similarity: 86.88% Mismatches: 24
Query Match: 84.10% Indels: 0
DB: 14 Gaps: 0
US-10-006-922a-12 (1-225) x AaA54911 (1-2241)
QY 1 MetArgSerSerLysAaValLleLysGluPheMetArgPheLysValArgMetGluGly 20
DB 805 ATGGCTCTCTCCAGAGACGTGATCAAGAGTTGATGCGCTTCAAGGTGGCGCATGGAGGC 864
QY 21 ThrValAaNGLYhIsgLupheGluileGluGlyGLYGLYGLYArgProTyrGluGly 40
DB 865 TCCGTGAACGGCCACGATTCGAGATCGAGGGCGAGGGCGCGCCCTTACGAGGGC 924
QY 41 HlSaenThrValLysLeuLysValThrLysGLYGLYProLeuProPheAlaTPAaPile 60
DB 925 ACCGAGCGCCCAAGCTGAAGGTGACCAAGGCGCGCCCTTCCCTTCCCTGGAGATC 984
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHlSaProAlaPilePro 80
DB 985 CTGTCCCTCAGTTCCAGTACGGCTCCAGAGCTTACGTGAAGCACCCCGCATCTCCC 1044
QY 81 AaPTrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAaPheGlu 100
DB 1045 GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGAACTTCGAG 1104
QY 101 AaPGLyGLYValValThrValThrGlnAaPSeSerSerleuGlnAaPGLyCySpheileTyr 120
DB 1105 GACGCGCGCGGTGGACCGTGAACCGAGACTCTCTCTGAGAGACGGCGAGTTCATCTAC 1164
QY 121 LysValLysPheileGlyValAaPheProSerAaPGLYProValMetGlnLysLysThr 140
DB 1165 AAGGTGAAGCTGCGCGCACCACTTCCCTCGAGCGCGCGTAATGCAAGAAAGACC 1224
QY 141 MetGLYTrpGlnAaSerThrGlnLysLeuTyrProArgAaPGLyValLleLysGLYGlu 160
DB 1225 ATGGGCTGGAGGCTTCCACCGAGCGGATGTACCCGAGGACGCGCGCTGAAGGGCGAG 1284
QY 161 lIeHlalyaLaleuLysLeuLysAaPGLYGLYhIstYrLeuValGluPheLysSerile 180
DB 1285 ATCAAGATGAGGCTGAAGCTGAAGAGACGGGCGCACTACACCGCGAGGTCAAGACACC 1344
QY 181 TyrMetAlalybLysPProValGlnLeuProGLYTrYTrYTrYValAaPSeLysLysAaP 200
DB 1345 TACATGCGCAAGAGCGCGGTGACCTGCGCGGCTTCAAGACCATCAAGAGCTGAGC 1404
QY 201 lIeThSerHlSaenGluAaPGLYTrThrileValGluGlnTyrGluAaTgThrGluGLYArg 220
DB 1405 ATCACTCTCCACAAAGAGACTACACCATGTGGAACATGACGAGCGCGCGAGGGCGCG 1464
QY 221 HIs 221
DB 1465 CAC 1467

RESULT 82
AaG59565
```

ID	AD059565	standard; DNA; 675 BP.
XX	AC	AD059565;
XX	DT	07-OCT-2004 (first entry)
XX	DE	Discoosoma monomerised red fluorescent protein mRFP1 DNA.
XX	KW	optically detectable signal; protein fragment complementation assay; PCA;
KW	KW	molecular interaction detection; biomolecular interaction detection;
KW	KW	fluorescent protein; multi-colour PCA; drug discovery; target validation;
KW	KW	high-throughput screening; high-content screening; pathway mapping;
KW	KW	drug mechanism-of-action study; biosensor; diagnostic;
KW	KW	red fluorescent protein; spectral shift; mRFP1; gene; ds.
XX	OS	Discoosoma.
XX	PH	Key
XX	FT	1.675
XX	FT	/tag= a
XX	FT	/product= "Discoosoma monomerised red fluorescent protein"
XX	PN	US200437528-A1.
XX	PD	15-JUL-2004.
XX	PF	01-DEC-2003; 2003US-00724178.
XX	PR	02-FEB-1998; 98US-00017412.
XX	PR	07-FEB-2000; 2000US-00499464.
XX	PR	24-MAY-2002; 2002US-00154758.
XX	PR	29-JAN-2003; 2003US-00353090.
XX	PR	09-APR-2003; 2003US-0461133P.
XX	PA	(MICH/) WATSON MICHNICK S W.
XX	PA	(MACD/) MACDONALD M L.
XX	PA	(LANE/) LAMERDIN J.
XX	PI	Watson Michnick SW, Macdonald ML, Lamerdin J;
XX	DR	WPI: 2004-533363/51.
XX	DR	P-PSDB, AD059566.
XX	PT	Composition useful in protein fragment complementation assays for drug
XX	PT	discovery and high-throughput screening, comprising complementary
XX	PT	fragments of protein or mutant protein, generating optically detectable
XX	PT	signal when associated.
XX	PS	Disclosure; SEQ ID NO 15; 34pp; English.
XX	XX	
XX	XX	The invention describes a composition (I) comprising complementary
XX	XX	fragments of a protein or mutant protein, where the fragments generate an
XX	XX	optically detectable signal when associated, and each of the mutant
XX	XX	protein fragments is fused to a separate molecule. Also described are:
XX	XX	protein fragment complementation assays (PCAs) for detection of molecular
XX	XX	interactions, involving reassembling separate fragments from an optically
XX	XX	detectable protein, and detecting the reassembly by units of
XX	XX	reconstitution of activity of the optically detectable protein, where the
XX	XX	reassembly of the fragments is operated by the interaction of molecular
XX	XX	domains fused to each fragment, and is independent of other molecular
XX	XX	processes; detecting biomolecular interaction, involving selecting an
XX	XX	appropriate optically detectable protein, effecting fragmentation of the
XX	XX	optically detectable protein such that the fragmentation results in
XX	XX	reversible loss of protein function, fusing or attaching fragments of the
XX	XX	optically detectable protein separately to other molecules, reassociating
XX	XX	the protein fragments through interactions of the molecules that are
XX	XX	fused or attached to the fragments, and detecting the resulting optical
XX	XX	signal; designing and engineering of PCAs based on fluorescent protein;
XX	XX	and a method and composition for the construction of multi-color PCAs.
XX	XX	(I) is useful in PCAs and other assays for drug discovery, target
XX	XX	validation, high-throughput screening, high-content screening, pathway
XX	XX	mapping, drug mechanism-of-action studies, biosensor and diagnostics.
XX	XX	(I) is useful for engineering different colour PCAs for a variety of

CC	applications in biology and biotechnology. This sequence encodes a
CC	Discooma monomethylated red fluorescent protein from which PCA fragments
CC	can be isolated.
XX	
XX	
Sequence	675 BP; 147 A; 229 C; 207 G; 92 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	1.03e-115
Score:	1015.00
Percent Similarity:	88.69%
Best Local Similarity:	86.43%
Query Match:	83.61%
DB:	12
US-10-006-922A-12 (1-225) x ADQ59565 (1-675)	
QY	1 MetatgSerSerlytAAsnValIlelysgIunPhemetaArgPhelysValArgMetglugly 20
DB	1 ATGGGCTCTCCGAGAGAGTATCAAGAGTTCAATGCGCTTCAAGGTGGCATGAGGGG 60
QY	21 ThrValaanglyVhIsgIunPhelugIlegluglygluglygluglyArgProtyrGlugly 40
DB	61 TCCGGAAACGGCCAGAGTTCAGATCGAGGCGGAGGGGAGGGCGGCCCTACGAGGGC 120
QY	41 HisaAntThrVallyleuLybValThrlysglyglIProleuProhealAtTPaaplie 60
DB	121 ACCGACGCCGCCAACCTTAAGGTGACCAAGGGCGGGCCCTGCTTCGCTGGGACATC 180
QY	61 LeuSerProglInPhelntyrGlyserlybValTyrVallyhIaProhlaaPlepro 80
DB	181 CTGTCCCTCACTTCACATACGGCTCCAAAGGCTACGTAAAGACACCCGCCGACATCCC 240
QY	81 AapTyrLylyLeuSerPheProgluglyPheLygTPTgluArgValMetAsnPhelgu 100
DB	241 GACTACTTGAAGCTGCTTCCCGAGGGCTTCAAGTGGAGGCGCTGATGAATTCGCA 300
QY	101 AapgllygllyValThrValThrGlnaPserSerLeuGlnaPgllyCyPheileTyr 120
DB	301 GACGGCGGGGTGTACCGTGAACCAAGACTCTCTCGAAGGACGAGATTCACTTAC 360
QY	121 LybVallyPheIlegllyValAsnPheProSerAapgllyProValMetGlnlybThr 140
DB	361 AAGGTGAAGCTCGCGGACCAACTTCCCTCGAAGGGCCGGTAAAGAGAAAGACC 420
QY	141 MetgllyTPTgluAlaserThrGluArgLeuTyrProArgAapgllyValleuLygIglu 160
DB	421 ATGGGCTGGAGGCTCCACAGCAGCGAATGTACCCGAGACAGCGCCCTGAAAGGGCGAG 480
QY	161 IleHlyblyAlaLeuLybLeuLybAapgllygllyHlybTyrleuValGluPheLybSerIle 180
DB	481 ATCAAGATGAGGCTGAAGCTGAAGAGCGCGGCACTTACGACGCCGAGGTCAAGACACC 540
QY	181 TyrMetAlaLybLybProValGlnleuProgllyTyrTyrTyrValaAspSerlybLeuAap 200
DB	541 TACATGGCCAGAGGCCCTGACGCTGCCCGGCGCTTACAGACCGACATCAAGGTGAC 600
QY	201 IleThrSerHlybAanglybAapTyrThrIleValGlnGlnTyrGluArgThrGluArg 220
DB	601 ATCACTCCCAACAGAGGACTACCAATCGTGAACAGTACAGAGCGCGCGAGGGCGCG 660
QY	221 His 221
DB	661 CAC 663
RESULT 83	
ADL46282	
ID	ADL46282 standard; DNA; 678 BP.
XX	
XX	ADL46282;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Discooma red fluorescent protein (Dsked) variant mRFP1.1 coding sequence

XX red fluorescent protein; DsRed; fluorescence; red wavelength;  
 KW oligomerization; tetramerization; immunoassay; hybridization assay; gene;  
 KM db.  
 XX  
 OS Discosoma sp.  
 XX  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 1..678  
 FT /\*tag= a  
 FT /product= "DsRed variant mRFP1.1 protein"  
 PN W02003086446-A1.  
 XX  
 XX 23-OCT-2003.  
 PD  
 XX  
 PF 09-APR-2003; 2003MO-US010879.  
 XX  
 PR 10-APR-2002; 2002US-00121258.  
 PR 29-JUL-2002; 2002US-00209208.  
 XX  
 XX (REGC ) UNIV CALIFORNIA.  
 XX  
 XX Tsien RY, Campbell RE, Baird GS;  
 PI WPI; 2003-845265/78.  
 DR P-PSDB; ADL46281.  
 XX  
 PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
 PT reduced propensity to oligomerize, and encoding polynucleotide, useful  
 PT in molecular biology, e.g. in immunoassays or in tracking protein  
 PT movement in cells.  
 XX  
 XX  
 XX Claim 63; SEQ ID NO 80; 166pp; English.  
 PS  
 CC The invention relates to a polynucleotide sequence encoding a Discosoma  
 CC red fluorescent protein (DsRed) variant having a reduced propensity to  
 CC oligomerize. The protein variant comprises one or more amino acid  
 CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
 CC sequence, where the substitutions result in reduced propensity of the  
 CC DsRed variant to form tetramers and where the variant displays detectable  
 CC fluorescence of at least one red wavelength. The composition and methods  
 CC are useful in producing red fluorescent proteins having reduced  
 CC propensity for oligomerization, especially tetramerization. The protein  
 CC may be used in molecular biology and in other scientific applications,  
 CC such as in immunoassays or hybridization assays, or in tracking the  
 CC movement of proteins in cells. This sequence corresponds to the DsRed  
 CC variant mRFP1.1 coding sequence. The protein contains a Q66W and a T147S  
 CC mutation.  
 CC  
 XX  
 SQ Sequence 678 BP; 148 A; 227 C; 208 G; 95 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,45e-115 Length: 678  
 Score: 1012.00 Matches: 190  
 Percent Similarity: 88.69% Conservative: 6  
 Best Local Similarity: 85.97% Mismatches: 25  
 Query Match: 83.36% Indels: 0  
 DB: 11 Gaps: 0  
 US-10-006-922a-12 (1-225) x ADL46282 (1-678)  
 QY 1 MetArgSerSerLysAsnValIleIleYsGluPheMetArgPheLysValArgMetGluGly 20  
 DB 1 ATGGCCTCTCTCCGAGACGTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGAGGC 60  
 QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGlyArgProTyrGluGly 40  
 DB 61 TCCGTGAACGGCCACAGAGTTCAGATTCAGAGCCAGAGGCGCGCCCTTACGAGGC 120  
 QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
 DB 121 ACCCAGACCGCGCAAGCTGAAGGTGACCAAGGCGCGCCCTTCCGCTTGGAATC 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
 DB 181 CTGTCCCTTCAGTTCATGTACGGGCTCCAGGCTTACGTGAAGACCCCGCGCATCTCC 240  
 QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
 DB 241 GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATCTTGAG 300  
 QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 DB 301 GACGGCGGCGGTGTGACCGGTGACCCAGACTCTCCCTGCGAGGACGCGCGATTCATCTAC 360  
 QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
 DB 361 AAGGTGAAGCTGGCGGACCAACTTCCCTCCGAGGCGCCGTTATGCAAGAGAGACC 420  
 QY 141 MetGlyTTrpGluLaserThrGluArgLysTyrProArgAspGlyValLeuLysGlyGlu 160  
 DB 421 ATGGGCTGGAGGCTCTCTCCGAGCGGATGTACCCGAGGACGCGCGCTGAAAGGCGAG 480  
 QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
 DB 481 ATCAAGATGAGCTGAAGCTGAAGACGCGGCGCACTACGACGCCGAGGTCAAGACACC 540  
 QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
 DB 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCGCTTCAAGACCGACATCAAGCTGGAC 600  
 QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
 DB 601 ATCACTCCCAACAAGACGACTACCATCTGTGAACAGTACGAGGCGCGCGAGCGCGC 660  
 QY 221 His 221  
 DB 661 CAC 663

Search completed: January 12, 2006, 12:55:37  
 Job time : 528 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 12, 2006, 12:26:48 ; Search time 3746 Seconds  
(without alignments)  
2810.223 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVYKEFMRFKRMKG.....EDYTYEYERTGRRHLLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DBVax1D  
-Q=/cgn2.1/USPTO.epool.p/US10006922/runat.10012006.162504.19128/app.query.fasta\_1.391  
-DB=EST -QFMT=fastap -SUFFIX=p2nminscr.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500  
-DOCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=69 -ALIGN=500 -MODE=LOCAL  
-OUTFMT=prc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922 @CGN 1.1 8010 @runat.10012006.162504.19128 -NCPU=6 -ICPU=3  
-NO MAP -LARGEOUTERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_est3:\*

4: gb\_hic:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_est7:\*

9: gb\_ges1:\*

10: gb\_ges2:\*

11: gb\_ges3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

No matches found

Search completed: January 12, 2006, 14:59:07  
Job time : 3747 secs

**THIS PAGE BLANK (USPTO)**



QY 181 TyMetAlaIysLysProValGlnLeuProGlyTyrTyrTyrValIAspSerLysLeuAsp 200  
| | | | |  
Db 541 TACATGGCAAGAGCCTGTCAGCTTACCAAGGAGTACCTACCTGTTGATCTCAAACTGGAT 600  
QY 201 IleThSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
| | | | |  
Db 601 ATAAACAAGCCACAAGAACTATACATCGTTGACAGATGTAAGAACCAAGAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
| | | | |  
Db 661 CACCATCTGTTCTT 675

## RESULT 2

US-09-866-538-11  
; Sequence 11, Application US/09866538  
; Patent No. 6852849  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSIEN, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530-2  
; CURRENT APPLICATION NUMBER: US/09/866,538  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: *Discochroma* sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (54)..(731)  
US-09-866-538-11

## Alignment Scores:

Pred. No.: 4,72e-160 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-006-922a-12 (1-225) x US-09-866-538-11 (1-859)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValIArgMetGluGly 20  
| | | | |  
Db 54 ATGAGGCTCTTCCAAAGATGTTATCAAGAGATTCAATGAGTTTAAAGTTTCCGATCGAAGGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
| | | | |  
Db 114 ACGGTCATAGGACGAGATTGAAATAGAAAGCCAAAGAGAGGAGGCCATACGAAGGC 173  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyLysProLeuProPheAlaTrpAspIle 60  
| | | | |  
Db 174 CACAAATACCGTAAGCTTAAGTAACCAAGGGGAGCCTTTGCCATTGGCTGGGAATTT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValLysValLysHisProAlaAspIlePro 80  
| | | | |  
Db 234 TTGTCAACACAAATTTCAGTATGGAAGCAAGGTATATGCAAGCACTTCGCCGACATACCA 293  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
| | | | |  
Db 294 GACTATATAAAGCTGTCATTTCTGAAAGATTTTAAATGGAAAGGTCATGAACCTTTGAA 353  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
| | | | |  
Db 354 GACGGTGGCGTCGTATCGTAACCAAGATTCAGATTGGACGAGATGGCTGTTTCATCTAC 413  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
| | | | |  
Db 414 AAGGTCAAGTTCAATGGCGTGAACCTTTCCTTCGATGAGACCTGTTATGCAAAAGAAAGACA 473

QY 141 MetGlyTrpGluIAspSerThrGluArgLysTyrProArgAspGlyValLeuLysGlyGlu 160  
| | | | |  
Db 474 ATGGGCTGGGAACCAAGCACTGAGCGTTGTATCTCTGTATGGCGGTGGAAAGAGAG 533  
QY 161 HisHisValAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
| | | | |  
Db 534 ATTCATTAAGGCTCTGAAGCTGAAGAGCGGTGTCATTAACCTAGTTGAATTCAAAGATTT 593  
QY 181 TyMetAlaIysLysProValGlnLeuProGlyTyrTyrTyrValIAspSerLysLeuAsp 200  
| | | | |  
Db 594 TACATGGCAAGAGCCTGTCAGCTTACCAAGGAGTACCTATGTGACTCCAAACTGGAT 653  
QY 201 IleThSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
| | | | |  
Db 654 ATAAACAAGCCACAAGAACTATACATCGTTGACAGATGTAAGAACCAAGAGGAGCGC 713  
QY 221 HisHisLeuPheLeu 225  
| | | | |  
Db 714 CACCATCTGTTCTT 728

## RESULT 3

US-09-865-291-11  
; Sequence 11, Application US/09865291  
; Patent No. 6900304  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSIEN, Roger  
; APPLICANT: TING, Alice  
; APPLICANT: ZHANG, Jin  
; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION  
; FILE REFERENCE: REGEN1550  
; CURRENT APPLICATION NUMBER: US/09/865,291  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: *Discochroma* sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (54)..(731)  
US-09-865-291-11

## Alignment Scores:

Pred. No.: 4,72e-160 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-006-922a-12 (1-225) x US-09-865-291-11 (1-859)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValIArgMetGluGly 20  
| | | | |  
Db 54 ATGAGGCTCTTCCAAAGATGTTATCAAGAGATTCAATGAGTTTAAAGTTTCCGATCGAAGGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
| | | | |  
Db 114 ACGGTCATAGGACGAGATTGAAATAGAAAGCCAAAGAGAGGAGGCCATACGAAGGC 173  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyLysProLeuProPheAlaTrpAspIle 60  
| | | | |  
Db 174 CACAAATACCGTAAGCTTAAGTAACCAAGGGGAGCCTTTGCCATTGGCTGGGAATTT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValLysValLysHisProAlaAspIlePro 80  
| | | | |  
Db 234 TTGTCAACACAAATTTCAGTATGGAAGCAAGGTATATGTCAGAACCTTCGCCGACATACCA 293  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
| | | | |  
Db 294 GACTATATAAAGCTGTCATTTCTGAAAGATTTTAAATGGAAAGGTCATGAACCTTTGAA 353



QY 101 AspGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 DB 354 GAGGTGGCGTCTGTTACTGTATACCCAGATTCAGTTGCGAGATGGCTGTTTCATCTAC 413  
 QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
 DB 414 AAGGTCAAGTTCAATGGCGTGAACCTTCTCCATGAGACCTGTTAGCAAAAGAACACA 473  
 QY 141 MetGlyTyrGlnLysSerThrGlnArgLeuTyrProArgAspGlyValLleuLysGlyLys 160  
 DB 474 ATGGGCTGGGAGGCCAGCACTGACCGCTTGTATCTCTGATGGCGCTTGAAAGAGAG 533  
 QY 161 IleIleLysAlaLeuLysLeuLysAspGlyValIleTyrLeuValGlnPheLysSerIle 180  
 DB 534 ATTCATAGAGCTCTGAAAGCTGAAAGACGGTGGTCATTACCTAGTTGAATCAAAAGTATT 593  
 QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
 DB 594 TACATGGCAAGAACCTGTGCACTACCAAGGTACTACTATGTGACTCCAAACTGGAT 653  
 QY 201 IleThrSerHisAsnGlnAspTyrThrIleValGlnGlnTyrGlnArgThrGlnGlyArg 220  
 DB 654 ATAAACAAGCCCAAGAGACTATACATCGTGAAGATGAAAGAACCGAGGAGCGC 713  
 QY 221 HisHisLeuPheLeu 225  
 DB 714 CACCATCTGTTCTT 728

## RESULT 4

US-10-152-296-1  
 ; Sequence 1, Application US/10152296  
 ; Patent No. 6723537  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beele, Beau  
 ; APPLICANT: Rigel Pharmaceuticals, Incorporated  
 ; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
 ; FILE REFERENCE: 021044-000110US  
 ; CURRENT APPLICATION NUMBER: US/10/152,296  
 ; PRIOR FILING DATE: 2002-12-10  
 ; PRIOR APPLICATION NUMBER: US 60/291,871  
 ; PRIOR FILING DATE: 2001-05-18  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 723  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:mammalian  
 ; OTHER INFORMATION: codon-optimized variant (DARED) of Discosoma sp.  
 ; OTHER INFORMATION: "red" red fluorescent protein (RFP)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(723)  
 ; OTHER INFORMATION: DARED  
 ; US-10-152-296-1

Alignment Scores:  
 Pred. No.: 1,316-159 Length: 723  
 Score: 1210.00 Matches: 224  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.56% Mismatches: 0  
 Query Match: 99.67% Indels: 0  
 DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-152-296-1 (1-723)

QY 1 MetArgSerSerLysAsnValIleLysGlnPheMetArgPheLysValArgMetGlnGly 20  
 DB 4 GTGGCGTCTCTCCAAAGACTCATCAAGAGATTCAATCGCTTCAAGGTGGCATGAGGCG 63  
 QY 21 ThrValAsnGlnHisGlnPheGlnIleGlnGlyGlnGlyArgProTyrGlnGly 40

DB 64 ACCGTGAACGCCACCAAGTTCCAGATCGAGGCGGAGGCGCGCCCTCAAGAGGCG 123  
 QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheIleATPAAPile 60  
 DB 124 CACAAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCGCTTCCGCTGGAGATC 183  
 QY 61 LeuSerProGlnIleGlnTyrGlnLysSerLysValTyrValLysHisProAlaAspIlePro 80  
 DB 184 CTGTCCCGCCAGTTCCAGTACGAGCTCCAGAGTGTACGTGAAGACCCCGCCGACATCCCC 243  
 QY 81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTyrGlnArgValMetAsnPheGln 100  
 DB 244 GACTCAAGAAAGCTGTCTTCCCGAGGCGCTTCAAGTGGAGCGCGATGAAGAACTTCGAG 303  
 QY 101 AspGlyValValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 DB 304 GACGCGCGGTGGTGAACCGTGAACCAAGACTCTCCCTGCGAGGACGCGCTGCTCATCTAC 363  
 QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
 DB 364 AAGGTGAAGTTCAATCGCGTGAACCTTCCCTCGACGCGCCCGTATGCAAGAGAACAC 423  
 QY 141 MetGlyTyrGlnLysSerThrGlnArgLeuTyrProArgAspGlyValLleuLysGlyLys 160  
 DB 424 ATGGGCTGGAGGCGCTTCAACGAGCGCTGTACCCCGCAGCGCGCTGAAAGGCGAG 483  
 QY 161 IleIleLysAlaLeuLysLeuLysAspGlyValIleTyrLeuValGlnPheLysSerIle 180  
 DB 484 ATCCAAAGGCCCTCGAAGCTGAAGACGGCGGCACATCTGAGTGAAGTCAAGATATC 543  
 QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
 DB 544 TACATGGCAAGAACCCGTGCACTGCGGCTGACTACTACTAGTGAATCCAAAGCTGAGC 603  
 QY 201 IleThrSerHisAsnGlnAspTyrThrIleValGlnGlnTyrGlnArgThrGlnGlyArg 220  
 DB 604 ATCACTCCCAACAGAGACTACATCATCTGAGAGCATGACGACGACGAGGCGCGC 663  
 QY 221 HisHisLeuPheLeu 225  
 DB 664 CACCACTGTTCTTG 678

## RESULT 5

US-10-001-189-45  
 ; Sequence 45, Application US/10001189  
 ; Patent No. 6962810  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FRASER JR., MALCOLM J.  
 ; APPLICANT: LI, XU  
 ; APPLICANT: BEAM, TERESA  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING  
 ; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION  
 ; FILE REFERENCE: 835910-92098  
 ; CURRENT APPLICATION NUMBER: US/10/001,189  
 ; PRIOR FILING DATE: 2001-10-30  
 ; PRIOR APPLICATION NUMBER: 60/244,984  
 ; PRIOR FILING DATE: 2000-11-01  
 ; PRIOR APPLICATION NUMBER: 60/244,677  
 ; PRIOR FILING DATE: 2000-10-31  
 ; NUMBER OF SEQ ID NOS: 70  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 45  
 ; LENGTH: 6964  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:  
 ; OTHER INFORMATION: pXFP3-DsRed-ori sequence  
 ; US-10-001-189-45

Alignment Scores:

Pred. No.: 4.87e-158 Length: 6984  
 Score: 1210.00 Matches: 224  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.56% Mismatches: 0  
 Query Match: 99.67% Indels: 0  
 DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-001-189-45 (1-6984)

```

QY      1 MetArgSerSerLysAenValIleLysGluPheMetArgPheLysValIArgMetGluLy 20
DB      2974 GTGGGCTCTCCAGAGAGCTCATCAGAGAGTTCATGCGCTTCAGAGTGCATGAGAGG 3033
QY      21 ThrValAenGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
DB      3034 ACCGTGAAGCGCCAGAGCTTCAGAGTCAGAGGCGAGGCGAGGCGCCCTTACAGAGGC 3093
QY      41 HisAenThrValLysLeuLysValThrLysGlyGlyProLeuProPheA1aTrpAspIle 60
DB      3094 CACAGACCGGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTCGCCCTTCGCTGGGACATC 3153
QY      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
DB      3154 CTGCCCCCAGTCCAGTCCAGGCTCCAGGTGACGTGAAGCACCCCGCGACATCCCC 3213
QY      81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAenPheGlu 100
DB      3214 GACTACAGAGAGCTGTCTCTCCCGAGGCTTCAGAGTGAAGCGCGTGAATGAATTCAG 3273
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      3274 GAGGCGGCGCGGTGACCGTGAACCAAGACTCCTCCGAGAGCGGCTGTTCTCATCTAC 3333
QY      121 LysValLysPheIleGlyValAenPheProSerAspGlyProValMetGlnLysLysThr 140
DB      3334 AAGGTGAAGTTCATCGGCGTGAACCTCCCTCCGACGCGCCCGTAAATGCAAGAGAAC 3393
QY      141 MetGlyTrpGluA1aSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
DB      3394 ATGGGCTGGAGGCTCTCAACGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGCGGAG 3453
QY      161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB      3454 ATCCAGAGGCGCTGAAGCTGAAGAGACGGGCGGCACTACCTGTGAGTTCAGTTCATC 3513
QY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
DB      3514 TACATGGCCAGAGAGCCCGTGCAGCTGCCGCTACTACTACGTGGACTCCAGAGCTGAC 3573
QY      201 IleThrSerHisAenGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB      3574 ATCACTCTCCAGAGAGACTACACATCGTGAAGCATGACGAGCGCACCGAGGCGCGC 3633
QY      221 HisHisLeuPheLeu 225
DB      3634 CACCACCTGTTCTCG 3648
  
```

Search completed: January 12, 2006, 15:02:20  
 Job time : 182 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 12, 2006, 12:46:57 ; Search time 792 Seconds  
(without alignments)  
2349.255 Million cell updates/sec

Title: US-10-006-922a-12

Perfect score: 1214

Sequence: 1 MRSSKNVKEFMFKVMEG.....EDYTVQYERTGRHHLFL 225

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 95

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlj  
-Q=/cgn2\_1/USPTO\_bpool/p/US10006922/runcat\_10012006.162505.19171/app\_query.fasta\_1.391  
-DB=Published\_Applications\_NA\_Main -QFMT=fastcap -SUFFIX=p2nmlnscr.rnpbm  
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1  
-MATRIX=blom62 -TRANS=human40.cdi -LIST=500 -DOCALIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=69 -ALIGN=500 -MODE=LOCAL -OUTFMT=pct -NORM=ext  
-HEAFILE=500 -MITLEN=0 -MAXLEN=200000000  
-USER=US10006922 @cgn2\_1\_1549 @runcat\_10012006.162505.19171 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORE=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA\_Main:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1214	100.0	678	3	US-09-967-772-6	Sequence 6, Appli
2	1214	100.0	678	5	US-10-006-922-11	Sequence 11, Appli
3	1214	100.0	678	5	US-10-081-864-7	Sequence 7, Appli
4	1214	100.0	678	5	US-10-121-258-2	Sequence 2, Appli
5	1214	100.0	678	6	US-10-315-920-1	Sequence 1, Appli
6	1214	100.0	678	6	US-10-312-067-3	Sequence 3, Appli
7	1214	100.0	678	6	US-10-335-517-6	Sequence 6, Appli

8	1214	100.0	678	6	US-10-334-288-6	Sequence 6, Appli
9	1214	100.0	678	7	US-10-311-030-5	Sequence 5, Appli
10	1214	100.0	678	8	US-10-656-029-21	Sequence 21, Appli
11	1214	100.0	678	9	US-10-505-486-27	Sequence 27, Appli
12	1214	100.0	678	9	US-10-844-064A-1	Sequence 1, Appli
13	1214	100.0	678	9	US-10-931-304-2	Sequence 2, Appli
14	1214	100.0	859	3	US-09-999-745-66	Sequence 66, Appli
15	1214	100.0	859	3	US-09-866-538-11	Sequence 11, Appli
16	1214	100.0	859	3	US-09-794-308-11	Sequence 11, Appli
17	1214	100.0	859	3	US-09-865-291-11	Sequence 11, Appli
18	1214	100.0	859	3	US-10-433-640-12	Sequence 11, Appli
19	1214	100.0	859	8	US-10-885-988-11	Sequence 11, Appli
20	1214	100.0	859	8	US-10-857-622-11	Sequence 11, Appli
21	1214	100.0	3311	3	US-09-797-4968-3	Sequence 3, Appli
22	1210	99.7	681	5	US-10-006-922-35	Sequence 35, Appli
23	1210	99.7	681	5	US-10-121-258-3	Sequence 3, Appli
24	1210	99.7	681	5	US-10-121-258-23	Sequence 23, Appli
25	1210	99.7	681	7	US-10-311-030-8	Sequence 8, Appli
26	1210	99.7	681	9	US-10-931-304-3	Sequence 3, Appli
27	1210	99.7	681	9	US-10-931-304-23	Sequence 23, Appli
28	1210	99.7	713	7	US-10-311-030-11	Sequence 11, Appli
29	1210	99.7	713	7	US-10-311-030-12	Sequence 12, Appli
30	1210	99.7	723	5	US-10-152-296-1	Sequence 1, Appli
31	1210	99.7	723	7	US-10-739-656-1	Sequence 51, Appli
32	1210	99.7	1638	5	US-10-214-932-51	Sequence 29, Appli
33	1210	99.7	1647	5	US-10-214-932-75	Sequence 16, Appli
34	1210	99.7	4692	6	US-10-161-403-29	Sequence 29, Appli
35	1210	99.7	4692	7	US-10-433-640-16	Sequence 16, Appli
36	1210	99.7	5436	9	US-11-006-076-29	Sequence 29, Appli
37	1210	99.7	6984	5	US-10-169-050-46	Sequence 46, Appli
38	1210	99.7	7910	8	US-10-001-189-45	Sequence 45, Appli
39	1210	99.7	9320	8	US-10-169-050-20	Sequence 20, Appli
40	1210	99.7	9658	8	US-10-471-065-20	Sequence 20, Appli
41	1210	99.7	9658	8	US-10-609-019-4	Sequence 4, Appli
42	1210	99.7	9678	8	US-10-609-019-3	Sequence 3, Appli
43	1210	99.7	10263	6	US-10-609-019-2	Sequence 2, Appli
44	1207	99.4	678	6	US-10-315-920-3	Sequence 3, Appli
45	1203	99.1	7508	7	US-10-742-828-4	Sequence 4, Appli
46	1198	98.7	681	5	US-10-006-922-37	Sequence 37, Appli
47	1196	98.5	666	7	US-10-332-733-22	Sequence 22, Appli
48	1196	98.5	711	7	US-10-314-936-1	Sequence 1, Appli
49	1196	98.5	711	7	US-10-314-936-3	Sequence 3, Appli
50	1196	98.5	711	10	US-11-021-014-1	Sequence 1, Appli
51	1196	98.5	711	10	US-11-021-014-3	Sequence 3, Appli
52	1191	98.1	678	5	US-10-006-922-36	Sequence 36, Appli
53	1191	98.1	678	5	US-10-081-864-14	Sequence 14, Appli
54	1191	98.1	678	6	US-10-315-920-5	Sequence 5, Appli
55	1186	97.7	675	5	US-10-006-922-38	Sequence 38, Appli
56	1186	97.7	675	5	US-10-081-864-13	Sequence 13, Appli
57	1186	97.7	747	8	US-10-785-862-10	Sequence 10, Appli
58	1186	97.7	1050	5	US-10-060-857-7	Sequence 7, Appli
59	1186	97.7	4200	9	US-10-894-949-9	Sequence 9, Appli
60	1186	97.7	4300	9	US-10-894-949-8	Sequence 8, Appli
61	1186	97.7	7495	7	US-10-742-828-5	Sequence 5, Appli
62	1186	97.7	9731	9	US-10-510-363-5	Sequence 5, Appli
63	1186	97.7	9782	9	US-10-510-363-6	Sequence 6, Appli
64	1184	97.5	898	5	US-10-006-922-45	Sequence 45, Appli
65	1160	95.6	678	5	US-10-121-258-5	Sequence 5, Appli
66	1160	95.6	678	5	US-10-931-304-5	Sequence 5, Appli
67	1160	95.6	704	9	US-10-844-064A-3	Sequence 3, Appli
68	1132	92.3	678	7	US-10-423-668A-10	Sequence 40, Appli
69	1132	92.3	681	5	US-10-121-258-7	Sequence 7, Appli
70	1121	92.3	681	9	US-10-931-304-7	Sequence 7, Appli
71	1121	92.3	1809	10	US-11-052-001-5	Sequence 5, Appli
72	1119	92.2	678	5	US-10-006-922-43	Sequence 43, Appli
73	1119	92.2	678	5	US-10-081-864-11	Sequence 11, Appli
74	1095	90.2	705	9	US-10-931-304-82	Sequence 82, Appli
75	1095	90.2	1431	9	US-10-931-304-107	Sequence 107, App
76	1085.5	89.4	876	5	US-10-006-922-17	Sequence 17, Appli
77	1085.5	89.4	876	5	US-10-161-403-39	Sequence 39, Appli
78	1085.5	89.4	876	10	US-11-006-076-39	Sequence 39, Appli
79	1021	84.1	675	7	US-10-724-178-15	Sequence 15, Appli
80	1021	84.1	678	5	US-10-121-258-9	Sequence 9, Appli

81 1021 84.1 678 9 US-10-931-304-9 Sequence 9, Appl  
82 1014 83.5 678 9 US-10-931-304-109 Sequence 109, App  
83 1012 83.4 678 9 US-10-931-304-80 Sequence 80, Appl  
84 1007 82.9 678 9 US-10-931-304-105 Sequence 105, App  
85 1002 82.5 678 9 US-10-931-304-101 Sequence 101, App  
86 994 81.9 711 9 US-10-931-304-86 Sequence 86, Appl  
87 992 81.7 711 9 US-10-931-304-90 Sequence 90, Appl  
88 989 81.5 711 9 US-10-931-304-103 Sequence 103, App  
89 988 81.4 711 9 US-10-931-304-84 Sequence 84, Appl  
90 987 81.3 711 9 US-10-931-304-93 Sequence 93, Appl  
91 986 81.2 708 9 US-10-931-304-95 Sequence 95, Appl  
92 973 80.1 711 9 US-10-931-304-97 Sequence 97, Appl  
93 973 80.1 711 9 US-10-931-304-99 Sequence 99, Appl  
94 960 79.1 711 9 US-10-931-304-88 Sequence 88, Appl  
95 848 69.9 549 7 US-10-724-178-1040 Sequence 1040, Ap

RESULT 1  
US-09-967-772-6  
Sequence 6, Application US/09967772  
Patent No. US20020164577A1  
GENERAL INFORMATION:  
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSJEN, Roger  
APPLICANT: GONZALEZ, Jesus  
TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
FILE REFERENCE: REGEN1290-5  
CURRENT APPLICATION NUMBER: US/09/967, 772  
CURRENT FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: US 09/459, 956  
PRIOR FILING DATE: 1999-12-13  
PRIOR APPLICATION NUMBER: US 08/765, 860  
PRIOR FILING DATE: 1996-12-19  
PRIOR APPLICATION NUMBER: PCT/ US96/09652  
PRIOR FILING DATE: 1996-06-06  
PRIOR APPLICATION NUMBER: US 08/481, 977  
PRIOR FILING DATE: 1995-06-07  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 6  
LENGTH: 678  
TYPE: DNA  
ORGANISM: *Discoecia* sp "red"  
US-09-967-772-6

Alignment Scores:  
Pred. No.: 6, 83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-006-922a-12 (1-225) x US-09-967-772-6 (1-678)

QY 1 MecArgSerSerIySaenValIlelyeGluPheMetArgPheIySaValArgMetGluGly 20  
DB 1 ATGAGGCTTCCTCCAGAAAGTATATCAAGAGTTCATGAGTTTAAGTTCACATGGAAGA 60  
QY 21 ThrValaAngIyHAgIuPheGluIleGluGlyGluGlyAAGPProIyrrGluGly 40  
DB 61 ACCGTCATAGGCGACGAGTTTGAATATGAAGGCGAAGAGAGGAGGCGCATACGAAGGC 120  
QY 41 HisAsnThrValIySaenValIlelyeGluGlyGluGlyAAGPProIyrrGluGly 60  
DB 121 CACATATCCGTAAAGCTTAAGTACCAAGGGGAGACTTTGGCATTTGGTGGATATT 180  
QY 61 LeuSerProGluPheGluIyrrGlySerIySaenValIyrrValIyShiSPRoAlaSPLePro 80  
DB 181 TTGTCAACACAAATTTGATGTGAAGCAAGGTATATGTCAAGCACCTCGCGACATACCA 240

QY 81 AspIyrrIySaenSerPheProGluGlyPheIyrrGluGlyValIleAsnPheGlu 100  
DB 241 GACTATAAAGAGTGTGATTTCTCTGAAGATTTAAATGGAAAGGATCAATTTGA 300  
QY 101 AspGlyValIySaenValIlelyeGluGlyGluGlyAAGPProIyrrGluGly 120  
DB 301 GACGTCGCGTCTGATCTGAACCCAGATTCAGATTCAGATTCAGATTCATCTAC 360  
QY 121 IySaenValIySaenPheProIyrrGluGlyGluGlyAAGPProIyrrGluGly 140  
DB 361 AAGGTAAAGTCTTGGCGGAACTTCTCTGATGACCTGTATGCAAAAGAGACA 420  
QY 141 MecGlyrrGluGlyAAGPProIyrrGluGlyGluGlyAAGPProIyrrGluGly 160  
DB 421 ATGGCTGGGAAGCCAGCATGAGCTTTGTATCCGTGATGCGGTGTTGAAGAGAG 480  
QY 161 IleHisIySaenValIySaenPheProIyrrGluGlyGluGlyAAGPProIyrrGluGly 180  
DB 481 ATTCATTAAGCTCTGAAGCTGAAGAGCGGTGATCTTACCTGATTAATCAAAAGTATT 540  
QY 181 TyrMetAlaIySaenPheProIyrrGluGlyGluGlyAAGPProIyrrGluGly 200  
DB 541 TACATGCAAAAGAGCTGTGACGTACCAAGGATCTACTATGTATGATCTCAAACTGAT 600  
QY 201 IleThrSerHisAsnGluAspIyrrThrIleValIleGluGlyGluGlyAAGPProIyrrGluGly 220  
DB 601 ATTAACAAGCCACACAGAAAGCTATACATCGTTGACAGATGAAAGAACCGAGGAGCG 660  
QY 221 HisIleuPheIeu 225  
DB 661 CACCATCTGTTCTCT 675

RESULT 2  
US-10-006-922-11  
Sequence 11, Application US/10006922  
Publication No. US20020197676A1  
GENERAL INFORMATION:  
APPLICANT: Lukyanov, Sergey A.  
APPLICANT: Fradkov, Arcady F.  
APPLICANT: Labas, Yulii A.  
APPLICANT: Matz, Mikhail V.  
TITLE OF INVENTION: No. US20020197676A1 Chromophores/fluorophores and  
FILE REFERENCE: Clon-035CIP  
CURRENT APPLICATION NUMBER: US/10/006, 922  
CURRENT FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: 09/120, 330  
PRIOR FILING DATE: 1998-12-11  
PRIOR APPLICATION NUMBER: 09/457, 898  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 09/458, 144  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 09/458, 477  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 09/457, 556  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 09/444, 338  
PRIOR FILING DATE: 1999-11-19  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 678  
TYPE: DNA  
ORGANISM: *Discoecia* species  
US-10-006-922-11

Alignment Scores:  
Pred. No.: 6, 83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0

DB: 5 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-006-922-11 (1-678)

QY 1 MetArgSerSerIysAsnValIleIleGluPheMetArgPheIysValArgMetGluGly 20

DB 1 ATGGAGCTCTCCCAAGAAAGTTATCAAGAGATTCAAGAGTTCAAGAGTTCCGATGGAAAGCA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyGluGlyGluGly 40

DB 61 ACCGTCAATGGGCGACGAGTTGGAAATAGAACGCAAGAGAGGGGAGGCCCATAGAGGC 120

QY 41 HisAsnThrValIysLeuIysValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60

DB 121 CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGGAGCTTGGCATTTGGTGAGATATT 180

QY 61 LeuSerProGlnPheGlnIleGlySerIysValIleValIysHisProAlaAspIlePro 80

DB 181 TTGTCAACCAAAATTCAGTATGAAAGCAAGGATATATGCAAGCAACCTGCCGACATACCA 240

QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100

DB 241 GACTATATAAAAGCTGTCTCTTCTGAAAGATTTAAATGGAAAGGTCATGCACTTTGAA 300

QY 101 AspGlyGlyValIleThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

DB 301 GAGGGTGGCGCTGTTACTGTAAACCAAGATTCCAGTTTGCAAGATGGCTGTTCATCTAC 360

QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIleIysThr 140

DB 361 AAGGTCAAGTTCATATGGGTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAAAGCA 420

QY 141 MetGlyTrpGluIleSerThrGluArgLeuTyrProArgAspGlyValIleIysGlyIle 160

DB 421 ATGGGCTGGGAAGCCAGCACTGAGGTTTGATCTCTGATGGCTTTGAAAGAGAG 480

QY 161 IleHisValAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180

DB 481 ATTCAATAGGCTCTGAACTGAAAGACGAGTGTCACTTCACTGAATTCATAAAGTATT 540

QY 181 TyrMetAlaIleIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200

DB 541 TACATGGCAAGAACCTGTGTGACCTACAGGGTACTCTAGTGTACTCCAAACTGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnIleTyrGluArgThrGluGlyArg 220

DB 601 ATACCAAGCCACAAAGAGACTATACATCTGTGACAGTATGAAGAAGAACGAGGAGCG 660

QY 221 HisHisLeuPheLeu 225

DB 661 CACCACTGTCTCCTT 675

RESULT 3

US-10-081-864-7

Sequence 7, Application US/10081864

Publication No. US2003002287A1

GENERAL INFORMATION:

APPLICANT: Lukanov, Sergey

APPLICANT: Lukanov, Konstantin

APPLICANT: Yanushkevich, Yuriy

APPLICANT: Savitsky, Alexander

APPLICANT: Fradkov, Arcady

TITLE OF INVENTION: No. US2003002287A1 Aggregating Fluorescent Proteins and

TITLE OF INVENTION: Methods for Using the Same

FILE REFERENCE: CLON-067

CURRENT APPLICATION NUMBER: US/10/081,864

CURRENT FILING DATE: 2002-06-19

PRIOR APPLICATION NUMBER: 10/006,922

PRIOR FILING DATE: 2001-12-04

PRIOR APPLICATION NUMBER: 60/270,983

PRIOR FILING DATE: 2001-02-21

NUMBER OF SEQ ID NOS: 30

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 678

TYPE: DNA

ORGANISM: Diacosoma sp

US-10-081-864-7

Alignment Scores:

Pred. No.: 6,93e-143 Length: 678

Score: 1214.00 Matches: 225

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 5 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-081-864-7 (1-678)

QY 1 MetArgSerSerIysAsnValIleIleGluPheMetArgPheIysValArgMetGluGly 20

DB 1 ATGGAGCTCTCCCAAGAAAGTTATCAAGAGATTCAAGAGTTCAAGAGTTCCGATGGAAAGCA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyGluGlyGluGly 40

DB 61 ACCGTCAACGGCGACGAGTTGAGATCAGAGGCGAGGCGAGGCGCCCTTACAGAGGC 120

QY 41 HisAsnThrValIysLeuIysValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60

DB 121 CACAAACCGTAAGCTTAAGGTAAACCAAGGGGGAGCTTGGCATTTGGTGAGATATT 180

QY 61 LeuSerProGlnPheGlnIleGlySerIysValIleValIysHisProAlaAspIlePro 80

DB 181 CTGTCCCCCGAGTTCAGATACGGCTCCAGGTGATACGTAAGCAACCCGCGCATCCCC 240

QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100

DB 241 GACTACAAAGAGCTGTCTCTCCGAGGGCTTCAAGTGGGAGCCGATGATGAATTCAG 300

QY 101 AspGlyGlyValIleThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

DB 301 GAGGCGGCGGTGGACCGTGAACCAAGACTCTCTCTGCAAGACGCTGCTTCACTAC 360

QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIleIysThr 140

DB 361 AAGGTCAAGTTCATCGGGTGAATCTCCCTCCGAGCGCCCGATGACAGAAAGAAC 420

QY 141 MetGlyTrpGluIleSerThrGluArgLeuTyrProArgAspGlyValIleIysGlyIle 160

DB 421 ATGGGCTGGAGGCTCCACCAAGGCGCTGTATACCCCGAGGCGGTGCTTAAGGGCGAG 480

QY 161 IleHisValAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180

DB 481 ATCCCAAGGCGCTGAACTGAAGGACGCGGCGCACTACCTGTGTGAGATTCAAGTCTAC 540

QY 181 TyrMetAlaIleIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200

DB 541 TACATGGCAAGAACCCGTGACGTGCCGCGTACTACTAGTGTACTCCAAAGTGAAC 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnIleTyrGluArgThrGluGlyArg 220

DB 601 ATCACTCCCAACAGAGACTACACATCTGTGAGCAGTACGAGCGACCGAGGGCGGC 660

QY 221 HisHisLeuPheLeu 225

DB 661 CACCACTGTCTCCTG 675

RESULT 4

US-10-121-258-2

Sequence 2, Application US/10121258

Publication No. US2003005983A1

GENERAL INFORMATION:

APPLICANT: Tsien, Roger

APPLICANT: Campbell, Robert

TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT

TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME

FILE REFERENCE: UC083.1CP2CP1  
CURRENT APPLICATION NUMBER: US/10/121.258  
CURRENT FILING DATE: 2002-04-10  
PRIORITY APPLICATION NUMBER: 09/794,308  
PRIORITY FILING DATE: 2001-02-26  
PRIORITY APPLICATION NUMBER: 09/866,538  
PRIORITY FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 678  
TYPE: DNA  
ORGANISM: *Discoecia* sp.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(678)  
OTHER INFORMATION: wild-type Dered  
US-10-121-258-2

Alignment Scores:  
Pred. No.: 6,83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-006-922a-12 (1-225) x US-10-121-258-2 (1-678)

1 MetArgSerSerLybAnValIleLygLuPheMetArgPheLybValArgMetGluGly 20  
1 ATGAGGCTCTTCCAGAAAGTTTATCAAGAGTTCATGAGTTTAAGTTTCCGATGGAAGA 60  
21 ThrValAnGlyYhiGluPheGluIleGluGlyGluGlyArgProTyArgGluGly 40  
61 ACCGTCAATGGGCGACGAGTTGAAATAGAAAGCGAAGAGGCGAGCGCATACGAAGGC 120  
41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTPAspIle 60  
121 CACAAATCCGTAAGCTTAAGTAACCAAGGGGGGACCTTGGCATTTGGTGGGATAATT 180  
61 LeuSerProGlnPheGlnTyArgLySerLybValTyValLybHisProAlaAspIlePro 80  
181 TTGTCAACACATTTCCATGATGGAAGCAAGGTATATGTCAAGCACCTGGCCGACATACCA 240  
81 AspTyrlLybLybLeuSerPheProGluGlyPheLybTyArgValMetAsnPheGlu 100  
241 GACTATATAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAGGTCATGAACTTTGAA 300  
101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTy 120  
301 GACGGTGGCGCTGTAAGTGAAGCCAGAGATTCAGTTTGGAGATGGCTGTTTCATCTAC 360  
121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLybThr 140  
361 AAGGTCAAGTTCAATGGCGTGAACCTTCTTCGATGGAAGCTGTTAAGCAAAAGACACA 420  
141 MetGlyTyArgGluAlaSerThrGluArgLeuTyArgArgAspGlyValLeuLybGlyGlu 160  
421 ATGGGCTGGGAAGCACACACTGACGCTTGTATCTGTGATGGCGGTGGAAAGGAGAG 480  
161 IleHisValAlaLeuLybLeuLybAspGlyGlyYhiGlyTyLeuValGluPheLybSerIle 180  
481 ATTCAATAAGGCTGAAGCTGAAGAGCGTGGTCAATTAAGTTGAATTCATAAAGTATT 540  
181 TyMetAlaLybLybProValGlnLeuProGlyTyTyTyTyValAspSerLybLeuAsp 200  
541 TACATGGCAAAAGAGCTGTGTCACTACCAAGGATACATATGTGACTCCAAACTGGAT 600  
201 IleThrSerHisAnGluAspTyArgThrIleValGluGlnTyArgValArgThrGluGlyArg 220  
601 ATAAACAAGCCACAACGAAGACTATACAAATCGTTGAGAGATGAAGAAACGAGGAGCGC 660

QY 221 HisHisLeuPheLeu 225  
DB 661 CACCAATCTGTCTCT 675

RESULT 5  
US-10-315-920-1

Sequence 1, Application US/10315920  
Publication No. US20030175809A1  
GENERAL INFORMATION:  
APPLICANT: Fradkov, Arcady Fedorovich  
APPLICANT: Terzikh, Alexey  
TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS  
TITLE OF INVENTION: FOR THEIR USE  
FILE REFERENCE: CLON-077CIP  
CURRENT APPLICATION NUMBER: US/10/315,920  
CURRENT FILING DATE: 2002-12-09  
PRIORITY APPLICATION NUMBER: 60/211,607  
PRIORITY FILING DATE: 2000-06-14  
PRIORITY APPLICATION NUMBER: PCT/US01/19097  
PRIORITY FILING DATE: 2001-06-13  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 678  
TYPE: DNA  
ORGANISM: *Artificial* Sequence  
OTHER INFORMATION: variant of sequence from *Discoecia* sp.  
US-10-315-920-1

Alignment Scores:  
Pred. No.: 6,83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-006-922a-12 (1-225) x US-10-315-920-1 (1-678)

1 MetArgSerSerLybAnValIleLygLuPheMetArgPheLybValArgMetGluGly 20  
1 ATGCGCTCTCCCAAGAACGTCATCAAGAGTTCATGCGCTTCAAGTGCACATGAGGCG 60  
21 ThrValAnGlyYhiGluPheGluIleGluGlyGluGlyArgProTyArgGluGly 40  
61 ACCGTGAACGCCACGAGTTCGAGATCGAGGCGAGGCGAGGCGCCCTTACGAGGCG 120  
41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTPAspIle 60  
121 CACAAACCGTGAAGCTGAAGTGAAGCCAGAGGCGGCGCTTGGCTTGGCTGGGACATC 180  
61 LeuSerProGlnPheGlnTyArgLySerLybValTyValLybHisProAlaAspIlePro 80  
181 CTGTCCCCCAAGTTCAGTACGCTCCAGGTGTACGTGAAGCACCCGCGCATCCCC 240  
81 AspTyrlLybLybLeuSerPheProGluGlyPheLybTyArgValMetAsnPheGlu 100  
241 GACTACAAAGAGGCTGTCTTCCCGAGGCTTCAAGGGGAGGCGCGTGAATCTTGGAG 300  
101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTy 120  
301 GACGGCGCGCTGTGAACCGTGAAGCCAGAGATCTCTCCCTGAGAGACGCTGTTCATCTAC 360  
121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLybThr 140  
361 AAGGTGAAGTTCAATGGCGTGAACCTTCCCTCCGACGCGCCCGTATGACGAAGAACCC 420  
141 MetGlyTyArgGluAlaSerThrGluArgLeuTyArgArgAspGlyValLeuLybGlyGlu 160  
421 ATGGGCTGGGAAGCTTCAACGAGCGCTGTACCCCGCACGCGCTGTGAAGGAGAG 480  
161 IleHisValAlaLeuLybLeuLybAspGlyGlyYhiGlyTyLeuValGluPheLybSerIle 180

Db ATCCAAAGGCGCTGAAGTGAAGACGGCCACTTACCTGGAGTTCAAGTCCATC 540  
Qy TyMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValLaaSerLysLeuAsp 200  
Db TACATGGCCAAAGAGCCGGTGCAGCTGCCGGCTACTACTACGCGACTCCAGCTGGAC 600  
Qy TlThSerHisAenGluAapTyrThrTlValGluGlnTyrGluAyrGthrGluGlyAyrG 220  
Db ATCACTCCCAACAAGAGACTACCATCTGTGAGAGTACGAGCGCACCGAGCGCGC 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCACTGTTCTTG 675

## RESULT 6

US-10-132-067-3  
Sequence 3, Application US/10132067  
Publication No. US2003020355A1  
GENERAL INFORMATION:  
APPLICANT: Bradbury, Andrew  
APPLICANT: Zeytun, Ahmet  
APPLICANT: Waldo, Geoffrey  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Fluorobodies: Binding Ligands with Intrinsic  
FILE REFERENCE: 021362-000600US  
CURRENT APPLICATION NUMBER: US/10/132,067  
CURRENT FILING DATE: 2002-04-24  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 678  
TYPE: DNA  
ORGANISM: *Discozyma* sp.  
FEATURE:  
FEATURE INFORMATION: red fluorescent protein (dared)  
NAME/KEY: CDS  
LOCATION: (1)..(678)  
OTHER INFORMATION: dared  
US-10-132-067-3

Alignment Scores:  
Pred. No.: 6,836-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-132-067-3 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGGTCTTCCAAAGATGTATCAAGAGTTCAAGAGTTTAAAGTTTCGCAATGGAAGA 60  
Qy ThrValAenGlyHisGluPheGluIleGluGluGluGluGluGluGluGluGluGluGlu 40  
Db 21 ACAGGCAATGGGCAAGAGTTTGAATTAAGAGCGAAGAGGAGGAGGCCATACGAAGCC 120  
Qy HisAenThrValLysLeuLysValThrLysGlyGlyProLeuProPheLaaTrpAspIle 60  
Db 41 CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGGAGCTTTGCCATTTGCTGGGATATT 180  
Qy LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProLaaAspIlePro 80  
Db 61 TTGTACACCAATTCATGATGAAAGCAAGTAAATGTCAACACCTTCGCCACATACCA 240  
Qy 81 AAPPYrLysLysLeuSerPheProGluGlyPheLysTrpGluAyrGValIleAenPheGlu 100  
Db 241 GACTATAAAGCTGTCAATTCCTGAAGATTTAAATGGAAAGGCTCATGAATTTGA 300

Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySpheIleTyr 120  
Db 301 GACGGTGGCCCTCGTACTGTAAACCCAGAGATTCAGTTTGCAGATGGCTGTTTCATCTAC 360  
Qy LysValLysPheIleGlyValAenPheProSerAspGlyProValMetGlnLysLeuThr 140  
Db 361 AAGGTCAAGTTCAATGGCCGGAAGCTTCTTCGATGAGACTGTTATGCAAAAGAGACA 420  
Qy 141 MetGlyTyrGluLaaSerThrGluArgLeuTyrProArgAspGlyValLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCAACAGACTAGAGCGTTGTATCTCGTATGCGCGTGTGAAGAGAG 480  
Qy 161 HisHisLysAlaLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLys 180  
Db 481 ATTCAATAGGCTCGAAGCTGAAGAGCGTGTATTAATTAATTAATTAATTAATTAATTA 540  
Qy 181 TyMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValLaaSerLysLeuAsp 200  
Db 541 TACATGGCCAAAGAGCCGTGCAAGCTACCAAGGTAATGTTGACTCCAAACTGGAT 600  
Qy 201 TlThSerHisAenGluAapTyrThrTlValGluGlnTyrGluAyrGthrGluGlyAyrG 220  
Db 601 ATACCAAGCCCAACAAGAGACTATACATCGTTGAGAGTATGAAGAACGAGGAGCGC 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCACTGTTCTTG 675

## RESULT 7

US-10-335-517-6  
Sequence 6, Application US/10335517  
Publication No. US20030207248A1  
GENERAL INFORMATION:  
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSJEN, Roger  
APPLICANT: GONZALEZ, Jesus  
TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
FILE REFERENCE: RESEN1290-5  
CURRENT APPLICATION NUMBER: US/10/335,517  
CURRENT FILING DATE: 2002-12-31  
PRIOR APPLICATION NUMBER: US/09/967,772  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: US 09/459,956  
PRIOR FILING DATE: 1999-12-13  
PRIOR APPLICATION NUMBER: US 08/765,860  
PRIOR FILING DATE: 1996-12-19  
PRIOR APPLICATION NUMBER: PCT/ US96/09652  
PRIOR FILING DATE: 1996-06-06  
PRIOR APPLICATION NUMBER: US 08/481,977  
PRIOR FILING DATE: 1995-06-07  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 6  
LENGTH: 678  
TYPE: DNA  
ORGANISM: *Discozyma* sp "red"  
US-10-335-517-6

Alignment Scores:  
Pred. No.: 6,836-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-335-517-6 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGGTCTTCCAAAGATGTATCAAGAGTTCAAGAGTTTAAAGTTTCGCAATGGAAGA 60  
Qy 21 ThrValAenGlyHisGluPheGluIleGluGluGluGluGluGluGluGluGluGluGlu 40

```

Db      61  ACCGTCATATGGGACCAAGTTGTAATATGAAAGCGGAGGAGGAGCCATACGAAAGC 120
Qy      41  HIsaAnthrValIysLeuIysValIthrLyegIyGlyProLeuProPhealATPaspIle 60
      121  CACAAATACCGTAAAGCTTAAGTTAACCAAGGGGGAGCCTTGGCATTGCTGGGATATT 180
Qy      61  LeuSerProGlnPheGlnTyrgIySerIysValItyrValIyHisProAlaApIlePro 80
      181  TTGTCAACCAATTTCAATGTGAAGCAAGGTATATGTCAAGACCTGCGGCATACCA 240
Qy      81  AspTyrIlybIysLeuSerPheProGluGlyPheIyTrpGluArgValIyMetAanPheGlu 100
      241  GACTATTAANAAGCTGCTATTTCTCGAAGATTTAAATGGGAAAGGCTCAAGACTTTGAA 300
Qy      101  AspGlyGlyValIyValIthrValIthrGlnAaspSerSerLeuGlnAaspGlyCySpheIleTyr 120
      301  GACGGTGGCGCTGCTACTGTTACCAAGATTCCAGTTGACAGATGGCTGTTTCATCTAC 360
Qy      121  LysValIyPheIleGlyValIaenPheProSerAaspGlyProValIyMetGlnIyLysThr 140
      361  AAGGTCAAGTTCATTTGGCGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAACACA 420
Qy      141  MetGlyTyrGluIyIaSerThrGluArgLeuTyrProArgAaspGlyValIleuLysGlyGlu 160
      421  ATGGCGTGGGAAAGCCAGCACTGACGCTTTGTATCTCTGATGGCGCTGTTGAAAGAGAG 480
Qy      161  ILeHisIyAlaIyLeuIySLeuIyAaspGlyGlyHisTyrLeuValIyGlnPheLysSerIle 180
      481  ATTCATTAAGGCTCTGAAGCTGAAGACGGTGGCTCATTAAGTTGAATTCAAAGATATT 540
Qy      181  TyrMetAlaIyLysPProValIyGlnLeuProGlyTyrTyrTyrValIyAaspSerIySLeuAasp 200
      541  TACATGGCAAAAGAGCTGTGACGACTACAGGGACTACTATGTTGACTCCTCAAACTGGAT 600
Qy      201  ILeThrSerHisAangIyAapTyrThrIleValIyGlnIyTyrgIyAaGthrGluIyArg 220
      601  ATTAACAAGCCACAAGAAAGCTATACATGCTTGAAGCAGTATGAAGAACAAGGAGGAGCGC 660
Db      221  HIsHisLeuPheLeu 225
      661  CACCATCTGTTCTT 675

```

RESULT 8  
US-10-334-288-6  
Sequence 6, Application US/10334288  
Publication No. US20040002123A1  
GENERAL INFORMATION:  
APPLICANT: THE REAGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSIEN, Roger  
TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
FILE REFERENCE: REGEN1290-5  
CURRENT FILING DATE: US/10/334,288  
PRIOR APPLICATION NUMBER: US/09/967,772  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: US 09/459,956  
PRIOR FILING DATE: 1999-12-13  
PRIOR APPLICATION NUMBER: US 08/765,860  
PRIOR FILING DATE: 1996-12-19  
PRIOR APPLICATION NUMBER: PCT/ US96/09652  
PRIOR FILING DATE: 1996-06-06  
PRIOR APPLICATION NUMBER: US 08/481,977  
PRIOR FILING DATE: 1995-06-07  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 6  
LENGTH: 678  
TYPE: DNA  
ORGANISM: *Diccosoma* sp "red"  
US-10-334-288-6

Alignment Scores:

Prod. No.:	6,83e-143	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922a-12 (1-225) x US-10-334-288-6 (1-678)

```

Qy      1  MetArgSerSerIyAanValIlyGlyGlnPheMetArgPheIySValIyArgMetGluGly 20
      1  ATGAGGCTTTCCAAAGATGTTATCAAGAGTTCATAGAGGTTTAAAGTTGCGATGGAAGGA 60
Qy      21  ThrValAangIyHisGlnPheGlnIleGlyGlnIyGlnIyGlnIyArgProTyrgIyGlu 40
      61  ACCGTCAATGGGACCAAGTTGTAATATGAAGGGAAGGAGGAGGAGGCCATACGAAAGC 120
Qy      41  HIsaAnthrValIysLeuIysValIthrLyegIyGlyProLeuProPhealATPaspIle 60
      121  CACAAATACCGTAAAGCTTAAGTTAACCAAGGGGGAGCCTTGGCATTGCTGGGATATT 180
Qy      61  LeuSerProGlnPheGlnTyrgIySerIysValItyrValIyHisProAlaApIlePro 80
      181  TTGTCAACCAATTTCAATGTGAAGCAAGGTATATGTCAAGACCTGCGGCATACCA 240
Qy      81  AspTyrIlybIysLeuSerPheProGluGlyPheIyTrpGluArgValIyMetAanPheGlu 100
      241  GACTATTAANAAGCTGCTATTTCTCGAAGATTTAAATGGGAAAGGCTCAAGACTTTGAA 300
Qy      101  AspGlyGlyValIyValIthrValIthrGlnAaspSerSerLeuGlnAaspGlyCySpheIleTyr 120
      301  GACGGTGGCGCTGCTACTGTTACCAAGATTCCAGTTGACAGATGGCTGTTTCATCTAC 360
Qy      121  LysValIyPheIleGlyValIaenPheProSerAaspGlyProValIyMetGlnIyLysThr 140
      361  AAGGTCAAGTTCATTTGGCGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAACACA 420
Qy      141  MetGlyTyrGluIyIaSerThrGluArgLeuTyrProArgAaspGlyValIleuLysGlyGlu 160
      421  ATGGCGTGGGAAAGCCAGCACTGACGCTTTGTATCTCTGATGGCGCTGTTGAAAGAGAG 480
Qy      161  ILeHisIyAlaIyLeuIySLeuIyAaspGlyGlyHisTyrLeuValIyGlnPheLysSerIle 180
      481  ATTCATTAAGGCTCTGAAGCTGAAGACGGTGGCTCATTAAGTTGAATTCAAAGATATT 540
Qy      181  TyrMetAlaIyLysPProValIyGlnLeuProGlyTyrTyrTyrValIyAaspSerIySLeuAasp 200
      541  TACATGGCAAAAGAGCTGTGACGACTACAGGGCTACTACTATGTTGACTCCTCAAACTGGAT 600
Qy      201  ILeThrSerHisAangIyAapTyrThrIleValIyGlnIyTyrgIyAaGthrGluIyArg 220
      601  ATTAACAAGCCACAAGAAAGCTATACATGCTTGAAGCAGTATGAAGAACAAGGAGGAGCGC 660
Db      221  HIsHisLeuPheLeu 225
      661  CACCATCTGTTCTT 675

```

RESULT 9  
US-10-311-030-5  
Sequence 5, Application US/10311030  
Publication No. US2004017107A1  
GENERAL INFORMATION:  
APPLICANT: Nelson, David  
APPLICANT: Zamiatra, Elize  
APPLICANT: Tsienn, Roger  
TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS  
FILE REFERENCE: 15916-032US1  
CURRENT FILING DATE: US/10/311,030  
PRIOR APPLICATION NUMBER: PCT/US01/04625  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: US 60/184,732



/ PRIOR FILING DATE: 2000-02-23  
 / NUMBER OF SEQ ID NOS: 12  
 / SOFTWARE: FASTSEQ for Windows Version 4.0  
 / SEQ ID NO 5  
 / LENGTH: 678  
 / TYPE: DNA  
 / ORGANISM: *Discozona sp.*  
 US-10-311-030-5

## Alignment Scores:

Pred. No.:	6,836-143	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-10-006-922a-12 (1-225) x US-10-311-030-5 (1-678)

```

QY 1 MetATGSeSerLyAsnValIleLySGluphMeATgPhelYsValArgMeGluGly 20
DB 1 ATGAGGCTCTCCAAAGATGTTATCAAGAGGTTCAAGGTTAAAGTTGCGATGAGGA 60
QY 21 ThrValaangLYHIGluphEgLuIleGluGlyGluGlyYArgProTYrGluGly 40
DB 61 ACGGTCATGGGCAAGATTGAAATAGAAAGCGAAGGAGGAGGCGCATACGAAGGC 120
QY 41 HIsAnThrValLyLeuLyValThrLySGlYrProLeuProPhelATrPaSpIle 60
DB 121 CACAATACCGTAAAGCTTAAGGTAACAAAGGGGAGCTTTGCCATTGCTTGGGATTT 180
QY 61 LeuSerProGlnPhEgIntYrGlySerLySValTYrValLyHIsProAlaSpIlePro 80
DB 181 TTGTCAACCAATTTCAAGATGGAAGCAAGTATATGTCAACACCCCTGCCACATACCA 240
QY 81 AApTYrLyLyLeuSerPhEProGluGlyPhelyeTPrGluArgValMeLsnPhEgLu 100
DB 241 GACTATTAAGAACTGTCATTTCTCGAAGATTTAAATGGAAAGGGTCATGAACCTTGA 300
QY 101 AApGlyGlyValValThrValThrGlnApsSerSerLeuGlnApsGlyCySpheIleTYr 120
DB 301 GACGGGCGGCTGTAAGCTTAACCCAGATTCAGATTTCAGAGATGGCTGTTCACTAC 360
QY 121 LySValLyPhEgIleGlyValaAnPhEProSeArSpGlyProValMeGlnLySlyThr 140
DB 361 AAGGTCATTAAGCTGTCATTTCTCGAAGATTTAAATGGAAAGGGTCATGAACCTTGA 420
QY 141 MeGlyTYrPgluaLaseThrGluArgLeuTYrProArgApsGlyValaLeuLySGly 160
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCTTTGTATCCCTCGATGGCGCTGTTGAAGAGAG 480
QY 161 ILeHIsLyAlaLeuLyLeuLySAspGlyGlyHIsTYrLeuValGluPhelySerIle 180
DB 481 ATTCAATTAAGCTCTGAAGCTGAAGACGGTGTGCATTAAGTTCAATTAATTAATTA 540
QY 181 TYrMeAlaLyLySProValGlnLeuProGlyTYrTYrTYrValaAspSerLySleuAap 200
DB 541 TACATGGCAAGAGGCTGTGCGCTACCAAGGTTACTATGTTGACTCCAAACCTGANT 600
QY 201 ILeThSerHIsAnGluApsTYrThrIleValGluGlnTYrGluArgThrGluGlyArg 220
DB 601 ATTAACAAGCACAAGAGCTATACATCGTTGACAGATGAAAGAACCGAGGAGCGC 660
QY 221 HIsHIsLeuPhelEu 225
DB 661 CACCATCTGTTCTT 675
  
```

RESULT 10  
 US-10-656-029-21  
 / Sequence 21, Application US/10656029  
 / Publication NO: US2005000367A1  
 / GENERAL INFORMATION:  
 / APPLICANT: VERTEX PHARMACEUTICALS INC.

/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF  
 / FILE REFERENCE: VPI/02-143WO2  
 / CURRENT APPLICATION NUMBER: US/10/656, 029  
 / PRIOR FILING DATE: 2003-09-05  
 / PRIOR APPLICATION NUMBER: 60/408,297  
 / PRIOR FILING DATE: 2002-09-05  
 / NUMBER OF SEQ ID NOS: 86  
 / SOFTWARE: PatentIn version 3.2  
 / SEQ ID NO 21  
 / LENGTH: 678  
 / TYPE: DNA  
 / ORGANISM: *Discozona sp.*  
 / FEATURE:  
 / NAME/KEY: CDS  
 / LOCATION: (1)..(678)  
 / OTHER INFORMATION: fluorescent protein  
 US-10-656-029-21

## Alignment Scores:

Pred. No.:	6,836-143	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-10-006-922a-12 (1-225) x US-10-656-029-21 (1-678)

```

QY 1 MetATGSeSerLyAsnValIleLySGluphMeATgPhelYsValArgMeGluGly 20
DB 1 ATGAGGCTCTCCAAAGATGTTATCAAGAGGTTCAAGGTTAAAGTTGCGATGAGGA 60
QY 21 ThrValaangLYHIGluphEgLuIleGluGlyGluGlyYArgProTYrGluGly 40
DB 61 ACGGTCATGGGCAAGATTGAAATAGAAAGCGAAGGAGGAGGCGCATACGAAGGC 120
QY 41 HIsAnThrValLyLeuLyValThrLySGlYrProLeuProPhelATrPaSpIle 60
DB 121 CACAATACCGTAAAGCTTAAGGTAACAAAGGGGAGCTTTGCCATTGCTTGGGATTT 180
QY 61 LeuSerProGlnPhEgIntYrGlySerLySValTYrValLyHIsProAlaSpIlePro 80
DB 181 TTGTCAACCAATTTCAAGATGGAAGCAAGTATATGTCAACACCCCTGCCACATACCA 240
QY 81 AApTYrLyLyLeuSerPhEProGluGlyPhelyeTPrGluArgValMeLsnPhEgLu 100
DB 241 GACTATTAAGAACTGTCATTTCTCGAAGATTTAAATGGAAAGGGTCATGAACCTTGA 300
QY 101 AApGlyGlyValValThrValThrGlnApsSerSerLeuGlnApsGlyCySpheIleTYr 120
DB 301 GACGGGCGGCTGTAAGCTTAACCCAGATTCAGATTTCAGAGATGGCTGTTCACTAC 360
QY 121 LySValLyPhEgIleGlyValaAnPhEProSeArSpGlyProValMeGlnLySlyThr 140
DB 361 AAGGTCATTAAGCTGTCATTTCTCGAAGATTTAAATGGAAAGGGTCATGAACCTTGA 420
QY 141 MeGlyTYrPgluaLaseThrGluArgLeuTYrProArgApsGlyValaLeuLySGly 160
DB 421 ATGGGCTGGGAAGCCAGCACTGAGGCTTTGTATCCCTCGATGGCGCTGTTGAAGAGAG 480
QY 161 ILeHIsLyAlaLeuLyLeuLySAspGlyGlyHIsTYrLeuValGluPhelySerIle 180
DB 481 ATTCAATTAAGCTCTGAAGCTGAAGACGGTGTGCATTAAGTTCAATTAATTAATTA 540
QY 181 TYrMeAlaLyLySProValGlnLeuProGlyTYrTYrTYrValaAspSerLySleuAap 200
DB 541 TACATGGCAAGAGGCTGTGCGCTACCAAGGTTACTATGTTGACTCCAAACCTGANT 600
QY 201 ILeThSerHIsAnGluApsTYrThrIleValGluGlnTYrGluArgThrGluGlyArg 220
DB 601 ATTAACAAGCACAAGAGCTATACATCGTTGACAGATGAAAGAACCGAGGAGCGC 660
  
```

Qy 221 HisHisteupheleu 225  
 Db 661 CACCATCTGTTCTT 675

## RESULT 11

US-10-505-486-27  
 ; Sequence 27, Application US/10505486  
 ; Publication No. US20050118639A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Takeda Chemical Industries, Ltd.  
 ; TITLE OF INVENTION: Determination of a ligand  
 ; FILE REFERENCE: P03-0006PCT  
 ; CURRENT APPLICATION NUMBER: US/10/505,486  
 ; CURRENT FILING DATE: 2004-08-20  
 ; PRIOR APPLICATION NUMBER: JP 2002-45728  
 ; PRIOR FILING DATE: 2002-02-22  
 ; PRIOR APPLICATION NUMBER: JP 2002-213949  
 ; PRIOR FILING DATE: 2002-07-23  
 ; PRIOR APPLICATION NUMBER: JP 2002-298237  
 ; NUMBER OF SEQ ID NOS: 233  
 ; SEQ ID NO 27  
 ; LENGTH: 678  
 ; TYPE: DNA  
 ; ORGANISM: *Discosoma* sp.  
 US-10-505-486-27

## Alignment Scores:

Pred. No.: 6,83e-143 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-505-486-27 (1-678)

Qy 1 MetArgSerSerlybAanValIlelyeGluPheMetArgPheLybValArgMetGluGly 20  
 Db 1 ATGAGGCTCTTCCAGAAATGTAATGAGAGTTCAAGAGTTTAAAGTTTAAAGTTTAAAGAG 60  
 Qy 21 ThrValAenGlyHsiGluPheGluIleGluGlyGluGlyArgProTyrgLugly 40  
 Db 61 ACGGTCAATGGGCGACGAGTTGAAATAGAACGCCAAGAGAGGGAGGCCATTACGAAGGC 120  
 Qy 41 HisAanThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60  
 Db 121 CACAATACCGTAAGCTTAAGTAAACCAAGGGGGGACCTTGGCATTTGGTGGGATATT 180  
 Qy 61 LeuSerProGluPheGluIleGlySerLybValTyrgLysbHsiAspAlaAspIlePro 80  
 Db 181 TTGTCCACACCAATTTCAAGTATGAAAGCAAGGATATGCAAGCACTTGGCCGACATACCA 240  
 Qy 81 AspTyrgLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100  
 Db 241 GACATTAATAAAGCTGTCAATTTCTGAAAGATTAAATGGAAAGGTCATGAACCTTGA 300  
 Qy 101 AspGlyGlyValValThrValThrGlnAspSerLeuGlnAspGlyCysPheIleTyrg 120  
 Db 301 GACGTGGCGTGGTAACTGTAAGTAAACCAAGGATTCAGTTTGGAGATGGCTGTTTCATCTAC 360  
 Qy 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLeuThr 140  
 Db 361 AAGGTCAAGTTCAATTTGCGTGAACCTTCTCCATGGAACCTGTTATGCAAAAGAGCA 420  
 Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrgProArgAspGlyValLeuLybGlyGlu 160  
 Db 421 ATGGGCTGGGAAGCAGACAGACAGAGGTTTGTATCTCTGTATGGCGTCTTAAAGAGAG 480  
 Qy 161 IleHsiLybAlaLeuLybLeuLybAspGlyGlyHsiTyrgLeuValGluPheLybSerIle 180  
 Db 481 ATCATTAAGGCTCGAAGCTGAAGACGAGTGTGATTCATTAGTTGAATCAAAAGATTAT 540

Qy 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrgTyrgValAspSerLybLeuAsp 200  
 Db 541 TACATGGCAAGAGAGCTGTGACGCTACAGGGTACACTATGTTGACTCCAAACTGGAT 600

Qy 201 IleThrSerHsiAenGlyAspTyrgThrIleValGluGlnTyrgLysbHsiGlyArg 220  
 Db 601 ATAAAGCCACCAAGAACGAACTATACAAATGTTGAGCAGTATGAAAGACCGAGGAGCGC 660

Qy 221 HisHisteupheleu 225  
 Db 661 CACCATCTGTTCTT 675

## RESULT 12

US-10-844-064A-1  
 ; Sequence 1, Application US/10844064A  
 ; Publication No. US20050149994A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gluck, Benjamin  
 ; APPLICANT: Bevis, Brooke  
 ; TITLE OF INVENTION: RAPIDLY MATURING FLUORESCENT PROTEINS AND METHODS FOR USING THE S  
 ; FILE REFERENCE: 092234-9006  
 ; CURRENT APPLICATION NUMBER: US/10/844,064A  
 ; CURRENT FILING DATE: 2004-05-11  
 ; PRIOR APPLICATION NUMBER: PCT/US02/40539  
 ; PRIOR FILING DATE: 2002-12-18  
 ; PRIOR APPLICATION NUMBER: US 60/341,723  
 ; PRIOR FILING DATE: 2001-12-19  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 678  
 ; TYPE: DNA  
 ; ORGANISM: *Discosoma*  
 US-10-844-064A-1

## Alignment Scores:

Pred. No.: 6,83e-143 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-844-064A-1 (1-678)

Qy 1 MetArgSerSerlybAanValIlelyeGluPheMetArgPheLybValArgMetGluGly 20  
 Db 1 ATGAGGCTCTTCCAGAAATGTAATGAGAGTTCAAGAGTTTAAAGTTTAAAGTTTAAAGAG 60  
 Qy 21 ThrValAenGlyHsiGluPheGluIleGluGlyGluGlyArgProTyrgLugly 40  
 Db 61 ACGGTCAATGGGCGACGAGTTGAAATAGAACGCCAAGAGAGGGAGGCCATTACGAAGGC 120  
 Qy 41 HisAanThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60  
 Db 121 CACAATACCGTAAGCTTAAGTAAACCAAGGGGGGACCTTGGCATTTGGTGGGATATT 180  
 Qy 61 LeuSerProGluPheGluIleGlySerLybValTyrgLysbHsiAspAlaAspIlePro 80  
 Db 181 TTGTCCACACCAATTTCAAGTATGAAAGCAAGGATATGCAAGCACTTGGCCGACATACCA 240  
 Qy 81 AspTyrgLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100  
 Db 241 GACATTAATAAAGCTGTCAATTTCTGAAAGATTAAATGGAAAGGTCATGAACCTTGA 300  
 Qy 101 AspGlyGlyValValThrValThrGlnAspSerLeuGlnAspGlyCysPheIleTyrg 120  
 Db 301 GACGTGGCGTGGTAACTGTAAGTAAACCAAGGATTCAGTTTGGAGATGGCTGTTTCATCTAC 360  
 Qy 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLeuThr 140  
 Db 361 AAGGTCAAGTTCAATTTGCGTGAACCTTCTCCATGGAACCTGTTATGCAAAAGAGCA 420

QY 141 MetGlyTTPGluAlaSerThrGluArgProGlyValIleuysgIygu 160  
 DB 421 ATGGGCTGGGAAGCCGACCTGAGCTTTGTATCCGTGATGGCTGTGAAAGAGAG 480  
 QY 161 TLeHleAlaIleuysIleuysIleuysApgIyGlyHleTyIleuValGluPheIysSerIle 180  
 DB 481 ATTCAATAGGCTGTGAAGCTGAAAGACGGTGTGATTAAGTTGAAATTCAAAAGTATT 540  
 QY 181 TTYMetAlaIleuysIleuysIleuysIleuysProGlyTyTyTyValAspSerIysIleuAsp 200  
 DB 541 TACATGGCAAGAGAGCTGTGACGCTACCAAGGCTACTATGTTGATTCGCAATCGGAT 600  
 QY 201 TLeThSerHisAngIuAspTyTyThIleValGluGlnTyTcIuArgThrGluGlyArg 220  
 DB 601 ATTAACAAGCCACAACGAAAGCTATACATCGTGAAGAGTATGAAAGAACCGAGGAGCGC 660  
 QY 221 HleHleuPheIeu 225  
 DB 661 CACCATCTGTCTT 675

## RESULT 13

US-10-931-304-2  
 / Sequence 2, Application US/10931304  
 / Publication No. US20050196768A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Tsien, Roger  
 / APPLICANT: Campbell, Robert  
 / APPLICANT: Baird, Geoffrey  
 / TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
 / TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
 / FILE REFERENCE: 39754-0831CP2CP3  
 / CURRENT APPLICATION NUMBER: US/10/931.304  
 / CURRENT FILING DATE: 2004-08-30  
 / PRIOR APPLICATION NUMBER: 10/209,208  
 / PRIOR FILING DATE: 2002-07-29  
 / PRIOR APPLICATION NUMBER: 10/121,258  
 / PRIOR FILING DATE: 2002-04-10  
 / PRIOR APPLICATION NUMBER: 09/866,538  
 / PRIOR FILING DATE: 2001-05-24  
 / PRIOR APPLICATION NUMBER: 09/794,308  
 / PRIOR FILING DATE: 2001-02-26  
 / NUMBER OF SEQ ID NOS: 110  
 / SOFTWARE: FastSeq for Windows Version 4.0  
 / SEQ ID NO 2  
 / LENGTH: 678  
 / TYPE: DNA  
 / ORGANISM: *Discofoma* sp.  
 / FEATURE:  
 / NAME/KEY: misc\_feature  
 / LOCATION: (1)..(678)  
 / OTHER INFORMATION: wild-type Dered  
 US-10-931-304-2

## Alignment Scores:

Pred. No.: 6,83e-143 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-2 (1-678)

QY 1 MetArgSerSerIysBaaValIleIysGluPheMetArgPheIysValArgMetGluGly 20  
 DB 1 ATGGAGTCTTCAGAAATGTTATCAAGAGTTTCATGAGTTTAAAGTTTCGATGGAAGA 60  
 QY 21 ThValAngIyHleGluPheGluIleGluGlyGluGlyArgProTyGluGly 40  
 DB 61 ACGGTCAATGGGACGAGTTTGAATAGAGCGAAGAGGAGGAGGCGCATACGAAGGC 120  
 QY 41 HleBaaThrValIyIleuysValThrIysGlyGlyProIeuProPheAlaTrpAspIle 60

DB 121 CAAATACCGTAAGCTTAAGGTAACCAAGGGGAGCCTTTCGCAATTGCGGATATT 180  
 QY 61 IeuSerProGlnPheGlnTyTcIySerIysValTyTyValIysHleProHlaAspIlePro 80  
 DB 181 TTYGTCAACCAATTCATGATGGAAGCAAGGTATATGTAAGACCTTCGACATACCA 240  
 QY 81 AApTyTyIleIysIleuSerPheProGluGlyPheIysTyTcIuArgValMetAsnPheGlu 100  
 DB 241 GACTATAAAGAGCTGTGATTCCTGTAAGGATTTAAATGGGAAAGGCTCATGAACCTTGA 300  
 QY 101 AApGlyIyValIyThrValThrGlnAspSerSerIleuGlnAspGlyCySphelIeTyR 120  
 DB 301 GACGGTGGCGTGTACTGTAACCCAGATTCAGATTGACAGATGCGTGTTCATCTTAC 360  
 QY 121 TyValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140  
 DB 361 AAGGTCAAGTTCATGGCGTGAACCTTTCCTCCAGAGACCTGTATGCAAAAGAAAGACA 420  
 QY 141 MetGlyTTPGluAlaSerThrGluArgProGlyValIleuysgIygu 160  
 DB 421 ATGGGCTGGGAAGCCGACCTGAGCTTTGTATCCTCGTATGGCGGTGAAAGAGAG 480  
 QY 161 TLeHleAlaIleuysIleuysIleuysApgIyGlyHleTyIleuValGluPheIysSerIle 180  
 DB 481 ATTCAATAGGCTGTGAAGCTGAAAGACGGTGTGATTAAGTTGAAATTCAAAAGTATT 540  
 QY 181 TTYMetAlaIleuysIleuysIleuysIleuysProGlyTyTyTyValAspSerIysIleuAsp 200  
 DB 541 TACATGGCAAGAGAGCTGTGACGCTACCAAGGCTACTATGTTGATTCGCAATCGGAT 600  
 QY 201 TLeThSerHisAngIuAspTyTyThIleValGluGlnTyTcIuArgThrGluGlyArg 220  
 DB 601 ATTAACAAGCCACAACGAAAGCTATACATCGTGAAGAGTATGAAAGAACCGAGGAGCGC 660  
 QY 221 HleHleuPheIeu 225  
 DB 661 CACCATCTGTCTT 675

## RESULT 14

US-09-999-745-66  
 / Sequence 66, Application US/09999745  
 / Patent No. US20020157120A1  
 / GENERAL INFORMATION:  
 / APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 / APPLICANT: Tsien, Roger Y.  
 / APPLICANT: Baird, Geoffrey  
 / TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS  
 / FILE REFERENCE: REGEN1470-1  
 / CURRENT APPLICATION NUMBER: US/09/999,745  
 / CURRENT FILING DATE: 2001-10-23  
 / PRIOR APPLICATION NUMBER: 09/316,920  
 / PRIOR FILING DATE: 1999-05-21  
 / NUMBER OF SEQ ID NOS: 67  
 / SOFTWARE: PatentIn version 3.0  
 / SEQ ID NO 66  
 / LENGTH: 859  
 / TYPE: DNA  
 / ORGANISM: *Discofoma* sp.  
 / FEATURE:  
 / NAME/KEY: CDS  
 / LOCATION: (54)..(731)  
 US-09-999-745-66

## Alignment Scores:

Pred. No.: 9.49e-143 Length: 859  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-09-999-745-66 (1-859)

```

QY      1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db      54 ATGAGGTCTTCCAAAGATGTTATCAAGAGGTTCAATAGGTTTAAGGTTGCGATGGAAGA 113
QY      21 ThrValaenglyHISgluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db      114 ACGGTCATAGGGCACAGGATTGAAATAGAAAGCGAAGAGAGGGGAGCCATACGAAGGC 173
QY      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60
Db      174 CACAATACCGTAAGCTTAAGGTAACCAAGGGGGAGCTTTGCGCATTTCTGGGATATT 233
QY      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db      234 TTGTACCAACAATTTCAAGTATGGAAGCAAGTATATGTCAAGACCCCTCCGACATACCA 293
QY      81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100
Db      294 GACTATTAATAAGCTGTCATTTCCCTGAAGGATTAAATGGAAAGGGTCAATGAACTTTGAA 353
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      354 GACGGTGGCGGTGTAATCTGAACCCAGATGCCAGTTCAGAGATGGCTGTTTCATCTAC 413
QY      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db      414 AAGGTCAAGTTCATTTGCGGCTGAACCTTTCTTCGATGACCTGTTATGCAAAAGAAACA 473
QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db      474 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCGTGAATGGCGTGTGAAGAGAGAG 533
QY      161 IleHisLysValAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db      534 ATTCATTAAGGCTCTGAAGCTGAAGCGGTGTCATTAAGTGAATTCAAAAGTATT 593
QY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200
Db      594 TACATGGCAAAAGAGCTGTGCAAGTACCAAGGTAATCTATGTTGACTCCAAACTGGAT 653
QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      654 ATTAACAAGCCACAAGAAAGACTATCAATCGTTGAGCAAGTATGAAGAAGAACGAGGAGCG 713
QY      221 HisHisLeuPheLeu 225
Db      714 CACCATCTGTCTCTT 728

RESULT 15
US-09-866-538-11
/ Sequence 11, Application US/09866538
/ Publication No. US20030032088A1
/ GENERAL INFORMATION:
/ APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
/ APPLICANT: TSIEN, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
/ FILE REFERENCE: REGEN1530-2
/ CURRENT APPLICATION NUMBER: US/09/866,538
/ CURRENT FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 29
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 11
/ LENGTH: 859
/ TYPE: DNA
/ ORGANISM: Discosoma sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (54) .. (731)
/ US-09-866-538-11
Alignment Scores:
Pred. No.: 9,49e-143 Length: 859

```

```

Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-10-006-922a-12 (1-225) x US-09-866-538-11 (1-859)

QY      1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db      54 ATGAGGTCTTCCAAAGATGTTATCAAGAGGTTCAATAGGTTTAAGGTTGCGATGGAAGA 113
QY      21 ThrValaenglyHISgluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db      114 ACGGTCATAGGGCACAGGATTGAAATAGAAAGCGAAGAGAGGGGAGCCATACGAAGGC 173
QY      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60
Db      174 CACAATACCGTAAGCTTAAGGTAACCAAGGGGGAGCTTTGCGCATTTGCTGGGATATT 233
QY      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db      234 TTGTACCAACAATTTCAAGTATGGAAGCAAGTATATGTCAAGACCCCTCCGACATACCA 293
QY      81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100
Db      294 GACTATTAATAAGCTGTCATTTCCCTGAAGGATTAAATGGAAAGGGTCAATGAACTTTGAA 353
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      354 GACGGTGGCGGTGTAATCTGAACCCAGATGCCAGTTCAGAGATGGCTGTTTCATCTAC 413
QY      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db      414 AAGGTCAAGTTCATTTGCGGCTGAACCTTTCTTCGATGACCTGTTATGCAAAAGAAACA 473
QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db      474 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCGTGAATGGCGTGTGAAGAGAGAG 533
QY      161 IleHisLysValAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db      534 ATTCATTAAGGCTCTGAAGCTGAAGCGGTGTCATTAAGTGAATTCAAAAGTATT 593
QY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp 200
Db      594 TACATGGCAAAAGAGCTGTGCAAGTACCAAGGTAATCTATGTTGACTCCAAACTGGAT 653
QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      654 ATTAACAAGCCACAAGAAAGACTATCAATCGTTGAGCAAGTATGAAGAAGAACGAGGAGCG 713
QY      221 HisHisLeuPheLeu 225
Db      714 CACCATCTGTCTCTT 728

RESULT 16
US-09-794-308-11
/ Sequence 11, Application US/09794308
/ Publication No. US20030170911A1
/ GENERAL INFORMATION:
/ APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
/ APPLICANT: TSIEN, Roger
/ APPLICANT: ZACHARIAS, David
/ APPLICANT: BAIRD, Geoffrey
/ TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
/ FILE REFERENCE: REGEN1530
/ CURRENT APPLICATION NUMBER: US/09/794,308
/ CURRENT FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 11
/ LENGTH: 859

```

```
/ TYPE: DNA
/ ORGANISM: Discosoma sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (54)..(731)
US-09-794-308-11
```

## Alignment Scores:

```
Pred. No.: 9,49e-143 Length: 859
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
```

US-10-006-922a-12 (1-225) x US-09-794-308-11 (1-859)

```
QY 1 MetArgSerSerlysaAnValIlelysgluPheMetArgPheIlyValArgMetGluGly 20
DB 54 ATGAGGCTCTTCCAAAGATGTTATCAAGAGTTGATGAGTTTAAAGTTCCGATGAAAGGA 113
QY 21 ThrValAsnGlyYHiegluPhegluilegluGlyGluGlyValArgProTyrgluGly 40
DB 114 ACGGTCAATGGGCAACAGTTTGAAATGAAAGCGAAGAGGGAGGCCATACGAAGGC 173
QY 41 HisAsnThrValIlyseulyValIThrlysglyProleuProPhealATPaapIle 60
DB 174 CACAAATACCTTAAGCTTAAGTACCAAGGGGGACCTTGGCAATTGCTGGGATATT 233
QY 61 LeuSerProGlnPheGlnTyrglySerlyValIlyValIlyshIspRoalaaapIlePro 80
DB 234 TTGTCAACCAATTTCAGTATGAGCAAGTATATGTCAAGCACCTGCCGACATACCA 293
QY 81 AspTyrllylyseulyseuPheProgluGlyPheIlystrpGluArgValMetAanPheGlu 100
DB 294 GACTATTAATAAGCTGTCATTTCTGAAGATTTAAATGGAAAGGTCATGAACCTTGAA 353
QY 101 AspGlyGlyValIlyThrValIThrGlnAspSerSerleuGlnaapGlyCyshelleTy 120
DB 354 GACGGTGGCGTCGTACTGTAACCAAGATTTCCAGTTGCAAGATGGCTGTTTCATCTAC 413
QY 121 LysValIlyshPheIleGlyValaAnPheProSerAspGlyProValMetGlnlyblyThr 140
DB 414 AAGGTCAAGTTCAATGGCGTGAACCTTCTCCATGACCTGTATGCAAAAGAAACA 473
QY 141 MetGlyTTrpGluIlyserThrGluArgleuTyProArgAspGlyValIleulysglyGlu 160
DB 474 ATGGGCTGGGAAGCAGCACTGACCGTTGTATCTCTGATGGCGCTGTTGAAGAGAG 533
QY 161 IleHieIlysaIaleulyseulyAspGlyGlyYHieTyrlleuValIglupheIySserIle 180
DB 534 ATTCATTAAGCTCTGAAGCTGAAAGACGGTGTCAATTAAGTTGAATCAAAAGATATT 593
QY 181 TytMetAlalylyspProValGlnleuProGlyTytyrTyryValaAspSerlyblyeAsp 200
DB 594 TACATGGCAAGAAGCTGTGACCTACCAAGGTACTATGTTGACTCCAAACTGAT 653
QY 201 IleThserHieAnGluuAspTyTrThrIleValIgluGlnTyrgluuArgThrGluGlyArg 220
DB 654 ATAAACAAGCCCAACGAAGACTATACATGTTGAGCATGTGAAGAACCGAAGGAGCGC 713
QY 221 HisHieIlyseuPheleu 225
DB 714 CACCATCTGTTCTT 728
```

## RESULT 17

```
US-09-865-291-11
/ Sequence 11, Application US/09865291
/ Publication No. US20030186229A1
/ GENERAL INFORMATION:
/ APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
/ APPLICANT: TSJEN, Roger
/ APPLICANT: TING, Alice
```

```
/ APPLICANT: ZHANG, Jin
/ TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION
/ FILE REFERENCE: REGEN1550
/ CURRENT APPLICATION NUMBER: US/09/865,291
/ CURRENT FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 42
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO: 11
/ LENGTH: 859
/ TYPE: DNA
/ ORGANISM: Discosoma sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (54)..(731)
US-09-865-291-11
```

## Alignment Scores:

```
Pred. No.: 9,49e-143 Length: 859
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
```

US-10-006-922a-12 (1-225) x US-09-865-291-11 (1-859)

```
QY 1 MetArgSerSerlysaAnValIlelysgluPheMetArgPheIlyValArgMetGluGly 20
DB 54 ATGAGGCTCTTCCAAAGATGTTATCAAGAGTTGATGAGTTTAAAGTTCCGATGAAAGGA 113
QY 21 ThrValAsnGlyYHiegluPhegluilegluGlyGluGlyValArgProTyrgluGly 40
DB 114 ACGGTCAATGGGCAACAGTTTGAAATGAAAGCGAAGAGGGAGGCCATACGAAGGC 173
QY 41 HisAsnThrValIlyseulyValIThrlysglyProleuProPhealATPaapIle 60
DB 174 CACAAATACCTTAAGCTTAAGTACCAAGGGGGACCTTGGCAATTGCTGGGATATT 233
QY 61 LeuSerProGlnPheGlnTyrglySerlyValIlyValIlyshIspRoalaaapIlePro 80
DB 234 TTGTCAACCAATTTCAGTATGAGCAAGTATATGTCAAGCACCTGCCGACATACCA 293
QY 81 AspTyrllylyseulyseuPheProgluGlyPheIlystrpGluArgValMetAanPheGlu 100
DB 294 GACTATTAATAAGCTGTCATTTCTGAAGATTTAAATGGAAAGGTCATGAACCTTGAA 353
QY 101 AspGlyGlyValIlyThrValIThrGlnAspSerSerleuGlnaapGlyCyshelleTy 120
DB 354 GACGGTGGCGTCGTACTGTAACCAAGATTTCCAGTTGCAAGATGGCTGTTTCATCTAC 413
QY 121 LysValIlyshPheIleGlyValaAnPheProSerAspGlyProValMetGlnlyblyThr 140
DB 414 AAGGTCAAGTTCAATGGCGTGAACCTTCTCCATGACCTGTATGCAAAAGAAACA 473
QY 141 MetGlyTTrpGluIlyserThrGluArgleuTyProArgAspGlyValIleulysglyGlu 160
DB 474 ATGGGCTGGGAAGCAGCACTGACCGTTGTATCTCTGATGGCGCTGTTGAAGAGAG 533
QY 161 IleHieIlysaIaleulyseulyAspGlyGlyYHieTyrlleuValIglupheIySserIle 180
DB 534 ATTCATTAAGCTCTGAAGCTGAAAGACGGTGTCAATTAAGTTGAATCAAAAGATATT 593
QY 181 TytMetAlalylyspProValGlnleuProGlyTytyrTyryValaAspSerlyblyeAsp 200
DB 594 TACATGGCAAGAAGCTGTGACCTACCAAGGTACTATGTTGACTCCAAACTGAT 653
QY 201 IleThserHieAnGluuAspTyTrThrIleValIgluGlnTyrgluuArgThrGluGlyArg 220
DB 654 ATAAACAAGCCCAACGAAGACTATACATGTTGAGCATGTGAAGAACCGAAGGAGCGC 713
QY 221 HisHieIlyseuPheleu 225
DB 714 CACCATCTGTTCTT 728
```

```

RESULT 18
US-10-433-640-12
/ Sequence 12, Application US/10433640
/ Publication No. US20040115792A1
/ GENERAL INFORMATION:
/ APPLICANT: Lichenberg-Frae, Hella
/ TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF
/ TITLE OF INVENTION: COMPLEX ENVIRONMENTAL CONTAMINATIONS
/ FILE REFERENCE: 1487/3
/ CURRENT APPLICATION NUMBER: US/10/433,640
/ CURRENT FILING DATE: 2003-10-02
/ PRIOR APPLICATION NUMBER: PCT/EP01/14610
/ PRIOR FILING DATE: 2001-12-12
/ PRIOR APPLICATION NUMBER: DE 10061872.3
/ PRIOR FILING DATE: 2000-12-12
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: Patentin version 3.2
/ SEQ ID NO 12
/ LENGTH: 859
/ TYPE: DNA
/ ORGANISM: Discosoma sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (54) .. (731)
US-10-433-640-12

Alignment Scores:
Pred. No.: 9,49e-143 Length: 859
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-433-640-12 (1-859)
QY 1 MetArGSeSerLyAsnValIleLySGluphMeArGpHeLyVaIArgMeGluGly 20
DB 54 ATGAGGTCCTCCAAAGATGTTATCAAGAGTTCATAGGTTTAAGGTTGCGATGGAAGCA 113
QY 21 ThrValaSnGlyHlaGluPhGluIleGluGlyGluGlyValArgProTyrgLugly 40
DB 114 ACGGTCAATGGGCAAGATTGAAATAGAAAGGGAAGAGGGAAGGCGCATACGAAGGC 173
QY 41 HlaSnThrValLyLeuLyValThrLySGlyGlyProLeuProPhaLatrpAaplle 60
DB 174 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGAACCTTGGCATTTGCGGATATT 233
QY 61 LeuSerProGluPhGluInTyrgLySerLyVaIlyValIlyHlaProIlaAapllePro 80
DB 234 TTGTCCACCAATTTAGTATGGAAGCAAGTATATGTCAAGACCTTCGCCACATACCA 293
QY 81 AspTyLyLyLeuSerPheProGluGlyPheLySTrGluArgValMeAanPheGlu 100
DB 294 GACTATAAAGAGTGTCAATTCCTGTAAGATTAAATGGGAAGGTCATGAACCTTTGA 353
QY 101 AapGlyGlyValValThrValThrGlnaPseSerLeuGlnaPseGlyCyapheIleTy 120
DB 354 GACGGTGGCGTGTCTGTATCCGATCCAGATTCAGGATGGCTGTTTCATCTAC 413
QY 121 LySValIlyPheIleGlyValaSnPheProSerAapGlyProValMeGlnIlyLySThr 140
DB 414 AAGGTCAACTCATTTGGCGTAACCTTCTCCGATGACCTGTATAGCAAAAGAAACA 473
QY 141 MetGlyTrGluAlaSerThrGluArgLeuTyProArgAapGlyValLeuLySGlyGlu 160
DB 474 ATGGGTGGGAAGCCGACATGAGCTTTGTATCCCTGTAAGCGCTGTGAAAGAGAG 533
QY 161 IleHlaLyValaLeuLyLeuLyAapGlyGlyHlaTyLeuValGluPhLySerIle 180
DB 534 ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGCTATTACCTAGTTGAATTAAGTAAT 593

```

```

QY 181 TyrMetAlaLyLyProValGlnLeuProGluTyTyTyTyValaAapSerLySLeuAap 200
DB 594 TACATGGCAAGAACCTGTGACGCTACCAAGGTACTACTATGTTGATCCAAACTGANT 653
QY 201 IleHlaSerHlaaSnGlyAapTyThrIleValGluGlnTyrgLuarGThrgLuglyArg 220
DB 654 ATAAACAACCAACAAAGAACTATCAATCGTTGAGCAGTATGAAGAAGCAAGGAGACGC 713
QY 221 HlaIleuPheLeu 225
DB 714 CACCATCTGTCTCT 728

RESULT 19
US-10-885-988-11
/ Sequence 11, Application US/10885988
/ Publication No. US20040259165A1
/ GENERAL INFORMATION:
/ APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
/ APPLICANT: TSJEN, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
/ FILE REFERENCE: REGEN1530-2
/ CURRENT APPLICATION NUMBER: US/10/885,988
/ CURRENT FILING DATE: 2004-07-06
/ PRIOR APPLICATION NUMBER: US/09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 29
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 11
/ LENGTH: 859
/ TYPE: DNA
/ ORGANISM: Discosoma sp.
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (54) .. (731)
US-10-885-988-11

Alignment Scores:
Pred. No.: 9,49e-143 Length: 859
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-885-988-11 (1-859)
QY 1 MetArGSeSerLyAsnValIleLySGluphMeArGpHeLyVaIArgMeGluGly 20
DB 54 ATGAGGTCCTCCAAAGATGTTATCAAGAGTTCATAGGTTTAAGGTTGCGATGGAAGCA 113
QY 21 ThrValaSnGlyHlaGluPhGluIleGluGlyGluGlyValArgProTyrgLugly 40
DB 114 ACGGTCAATGGGCAAGATTGAAATAGAAAGGGAAGAGGGAAGGCGCATACGAAGGC 173
QY 41 HlaSnThrValLyLeuLyValThrLySGlyGlyProLeuProPhaLatrpAaplle 60
DB 174 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGAACCTTGGCATTTGCGGATATT 233
QY 61 LeuSerProGluPhGluInTyrgLySerLyVaIlyValIlyHlaProIlaAapllePro 80
DB 234 TTGTCCACCAATTTCAGTATGGAAGCAAGTATATGTCAAGACCTTCGCCACATACCA 293
QY 81 AspTyLyLyLeuSerPheProGluGlyPheLySTrGluArgValMeAanPheGlu 100
DB 294 GACTATAAAGAGTGTCAATTCCTGTAAGATTAAATGGGAAGGTCATGAACCTTTGA 353
QY 101 AapGlyGlyValValThrValThrGlnaPseSerLeuGlnaPseGlyCyapheIleTy 120
DB 354 GACGGTGGCGTGTATCTGTAAACCAAGATTCAGTTGCAAGATGGCTGTTTCATCTAC 413
QY 121 LySValIlyPheIleGlyValaSnPheProSerAapGlyProValMeGlnIlyLySThr 140

```

DB 414 MAGTCAGTTCATGGCGGTGAACCTTCTCCGATGACCTGTATGCAAGAAAGACA 473  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuysgIygu 160  
DB 474 ATGGGCTGGGAAAGCCGACCTGAGCTTTGTATCCCTGATGGCGCTGTGAAGAGAG 533  
QY 161 HleHlAlaLeuAlaLeuAlaLeuysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 534 ATTCATTAAGCTCTGAAGCTGAAGAGCGGTGTCATTACCTAGTTGAATTCAAAGTATT 593  
QY 181 TyrMetAlaLysLeuProValGluLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 594 TACATGGCAAGAGAGCTGTGACGCTGACCTACTATGTGACTCTCAACTGGAT 653  
QY 201 HleHlSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 654 ATACAAAGCCACAAAGAGCTATCAATGCTTGACAGATGAAGAAAGAGGAGACGC 713  
QY 221 HleHlAlaLeuPheLeu 225  
DB 714 CACCATCTGTCTCT 728

## RESULT 20

US-10-857-622-11  
Sequence 11, Application US/10857622  
Publication No. US20050026234A1  
GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: VIOLIN, Jonathan  
APPLICANT: NEWTON, Alexandra  
APPLICANT: TSJEN, Roger  
APPLICANT: ZHANG, Jin  
TITLE OF INVENTION: Emission Ratiometric Indicators of Phosphorylation By C-Kinase  
FILE REFERENCE: 39754-0891 CPCICP2  
CURRENT FILING DATE: 2004-05-28  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: US 09/865,291  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: US 09/396,003  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: US 08/792,553  
PRIOR FILING DATE: 1997-01-31  
PRIOR APPLICATION NUMBER: US 594,575  
PRIOR FILING DATE: 1996-01-31  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 859  
TYPE: DNA  
ORGANISM: *Discozyma* sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54)..(731)  
US-10-857-622-11

## Alignment Scores:

Pred. No.: 9,496-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-857-622-11 (1-859)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 54 ATGAGGCTTCACAAAGATGTATCAAGAGTTCATGAGTTTAAGCTTCGATGAAAGGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGlyGlyArgProTyrGluGly 40  
DB 114 ACGGTCAATGGGACGAGTTTGAATATGAAGCGCAAGAGAGGAGGAGCCATACGAAGGC 173

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 174 CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGAGACCTTCCATTCTCGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 234 TTGTACCAACATTTTCAGTATGGAAGCAAGATATGTCAAGCACCTCTCCGACATACCA 293  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 294 GACTATTAAGAACTGTCAATTCCTGAAGATTTAAATGGAAAGGTCATGAACCTTTGA 353  
QY 101 AspGlyLysValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 354 GACGGGCGGTGTACTATACCAAGATTCAGATTCCAGATGGCTGTTCATCTAC 413  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 414 MAGTCAGTTCATTTGGCGGTGAACCTTCTCCGATGACCTGTATGCAAGAAAGACA 473  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuysgIygu 160  
DB 474 ATGGGCTGGGAAAGCCAGACCTGAGCTTTGTATCCCTGATGGCGCTGTGAAGAGAG 533  
QY 161 HleHlAlaLeuAlaLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 534 ATTCATTAAGCTCTGAAGCTGAAGAGCGGTGTCATTACCTTGAATTCAAAGTATT 593  
QY 181 TyrMetAlaLysLeuProValGluLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 594 TACATGGCAAGAGAGCTGTGACGCTGACGACAGGATCTACTATGTGATCTCAACTGGAT 653  
QY 201 HleHlSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 654 ATACAAAGCCACAAAGAGCTATCAATGCTTGACAGATGAAGAAAGAGGAGACGC 713  
QY 221 HleHlAlaLeuPheLeu 225  
DB 714 CACCATCTGTCTCT 728

## RESULT 21

US-09-797-496B-3  
Sequence 3, Application US/09797496B  
Publication No. US20030049597A1  
GENERAL INFORMATION:  
APPLICANT: Simon, Sanford M.  
APPLICANT: Chen, Yu  
TITLE OF INVENTION: Chimeric Fluorescent Enzymes and Uses Thereof  
FILE REFERENCE: 600-1-267  
CURRENT FILING DATE: 2002-05-24  
PRIOR FILING DATE: 2002-05-24  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 3311  
TYPE: DNA  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: *Discozyma* red fluorescent protein modified as described in specific  
OTHER INFORMATION: ication.  
US-09-797-496B-3

## Alignment Scores:

Pred. No.: 6,226-142 Length: 3311  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-09-797-496B-3 (1-3311)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

```

Db      289 ATGAGGTCTCCAGAAATGTTATCAAGAGGTTTCATGAGGTTTAAGGTTGCGATGAGAGA 348
Qy      21  ThrValAsnGlyYHsGluPheGluIleGluGluGluGluGluYArgProTyrGluGly 40
Db      349 ACCGTCAATGGGCAACAGATTGTAATAAGAAAGGGAAGAGGAGGCGCATACGAAGCC 408
Qy      41  HisAsnThrValIleLeuLeuValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60
Db      409 CACAAATACCGTAAAGCTTAAGGTACCAAGGGGGACCTTTGCCATTTCCTTGGGATATT 468
Qy      61  LeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHisProAlaAspIlePro 80
Db      469 TTGTACCAACAATTCATGATGGAACCAAGTATATGTCAGAGACCCTCCGACATACCA 528
Qy      81  AspTyrIleValLeuSerPheProGluGlyPheIleTyrGluArgValMetAsnPheGlu 100
Db      529 GACTATAAAGGTGTCAATTCCTGAAAGATTTAAATGGAAAGGCTCATGAACTTTGAA 588
Qy      101 AaPGIYGLVValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120
Db      589 GACGGTGGGCTGTTACTGTAACCCAGATTCCAGTTTGAGAGATGGCTGTTTCATCTAC 648
Qy      121 LysValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnIleYsThr 140
Db      649 AAGGTCAAGTTCATTGGCGTGAACCTTCTTCGATGGAAGCTGTTATGCAAAAGAAACA 708
Qy      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIleGlyGlu 160
Db      709 ATGGGCTGGGAAGCCGACCTGAGCTGTGTATCCCTGTAAGCGCTGTGTAAGAGAGAG 768
Qy      161 IleHisIleValLeuLeuLeuAspGlyValIleTyrLeuValGluPheIleSerIle 180
Db      769 ATTCATAAAGCTCTGAAGCTGAAGACGGAGGTCATTACCTAGTGAATTCAAAGTATT 828
Qy      181 TyrMetAlaIleLeuLeuLeuAspGlyValIleTyrTyrTyrValAspSerIleAsp 200
Db      829 TACATGGCAAGAGAGCTGTGACGTAACGAGGTACTAATGTTGACCTCAAACTGAT 888
Qy      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      889 ATTAACAAGGACACAAACGAACTAATCATCTGTGACAGTATGAAGAACGAGGGAGCC 948
Qy      221 HisHisIleuPheLeu 225
Db      949 CACCACTGTCTCTT 963

```

## RESULT 22

```

US-10-006-922-35
; Sequence 35, Application US/10006922
; GENERAL INFORMATION:
; APPLICANT: Lukanov, Sergey A.
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Terekikh, Alexey
; TITLE OF INVENTION: No. US2002019767A1el Chromophores/fluorophores and
; FILE REFERENCE: ClON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338

```

```

; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Discosoma species
US-10-006-922-35
Alignment Scores:
Pred. No.: 2,2e-142 Length: 681
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 5 Gaps: 0
US-10-006-922A-12 (1-225) x US-10-006-922-35 (1-681)
Qy      1  MetArgSerSerLysAsnValIleLeuGluPheMetArgPheLysValArgMetGluGly 20
Db      4  GTGCCCTCTCCAAAGAGCTCATCAAGAGATTGATGCGCTTCAAGGTGCGCATGAGAGGC 63
Qy      21  ThrValAsnGlyYHsGluPheGluIleGluGluGluGluYArgProTyrGluGly 40
Db      64  ACCGTGAACGGCCACAGAGTTCAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGC 123
Qy      41  HisAsnThrValIleLeuLeuValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60
Db      124  CACAAACACCGTGAACCTGAAGGTGACCAAGGGCGGCCCCCTTGCCTCGGACATC 183
Qy      61  LeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHisProAlaAspIlePro 80
Db      184  CTGTCCCCCACTTCAAGTACGCTCCAAAGGTGACGTAAGACACCCGCCACATCCCC 243
Qy      81  AspTyrIleValLeuSerPheProGluGlyPheIleTyrGluArgValMetAsnPheGlu 100
Db      244  GACCTCAAGAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAATCTTCGAG 303
Qy      101  AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120
Db      304  GACGGCGCGGTGGTACCGTGAACCCAGACCTCTCCGACGAGCGGCTGTAATGCAAGAGACC 363
Qy      121  LysValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnIleYsThr 140
Db      364  AAGGTGAAGTTCATGGCGTGAATCTTCCCTCCAGCGGCCCTGTAATGCAAGAGAGACC 423
Qy      141  MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIleGlyGlu 160
Db      424  ATGGCTGGAGGGCTCCACCGAGGCGCTGTACCCCGGACGGCGTGTGAAGGGCGAG 483
Qy      161  IleHisIleValLeuLeuLeuAspGlyValIleTyrLeuValGluPheIleSerIle 180
Db      484  ATCCACAAAGCCCTGAAGCTGAAGGACGGCGCCACTACCTGTGTGAGATTCAAGTCCATC 543
Qy      181  TyrMetAlaIleValProValGlnLeuProGlyTyrTyrTyrValAspSerIleAsp 200
Db      544  TACATGGCCAAAGAGCCCGTGAAGTGGCTGCCCGCTACTACGTGAGATCCAAAGCTGAGC 603
Qy      201  IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      604  ATCACTCCCAACAAGAGAGTACACCATCGTGAAGAGTACGAGCGACGAGCGAGGGCGGC 663
Qy      221  HisHisIleuPheLeu 225
Db      664  CACCACTGTCTCTT 678

```

## RESULT 23

```

US-10-121-258-3
; Sequence 3, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Teien, Roger

```



```
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ FILE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: UC083.LCP2CPI
/ CURRENT APPLICATION NUMBER: US/10/121,258
/ CURRENT FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 681
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: nucleotide sequence encoding Dered with mammalian
/ OTHER INFORMATION: codon usage
US-10-121-258-3
```

## Alignment Scores:

```
Prod. No.: 2,2e-142 Length: 681
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: Gaps: 0
```

US-10-006-922a-12 (1-225) x US-10-121-258-3 (1-681)

```
QY 1 MetArgSerSerlybAnValIlelybGluPheMetArgPheValArgMetGluGly 20
DB 4 GTCGGCTCTCCCAAGAACGTCAACAGAGTTTCATGGCTTCAAGGCGCATGAGGCG 63
QY 21 ThrValAsnGlyVHlsgLupheGluIleGluGlyGluGlyArgProTyrgLugly 40
DB 64 ACCGTAAAGCCCAAGATCGAGATCGAGGCGAGGCGCGCCCTTACGAGGCG 123
QY 41 HisAsnThrValIleuLeuValThrlybGlyGlyProLeuProPheAlaTPAspIle 60
DB 124 CACAAACCGGTGAAGCTGAAGTGAACCAAGGCGGCCCCCTGCCCCCTTGGGACATC 183
QY 61 LeuSerProGlnPheGlnTyrglySerlybValTyrglybHlspProAlaAspIlePro 80
DB 184 CTGTCCCCCAAGTTCAGATCGAGCTCCAGATGATGAAGACCCCGCGACATCCCC 243
QY 81 AspTyrllyblyLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100
DB 244 GACTACAAAGAGCTGTCTCCCGAGGCGCTTCAAGTGGAGCGCGCTGATGAACCTTCGAG 303
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerleuGlnAspGlyCysPheIleTy 120
DB 304 GACGGCGCGCTGTGATGACCTGAACCAAGACTCTCCCTGCAAGGCGGCTCTTCATTTAC 363
QY 121 LysValIlePheIleGlyValIleAsnPheProSerAspGlyProValMetGlnlyblybThr 140
DB 364 AAGGTGAAGTTCACTGGCGCTGAACCTTCCCTCGAAGCGGCCCGTAAGCAAGAAACACC 423
QY 141 MetGlyTrpGluIleAsnSerPheProGluArgLeuTyrglybTrpGluArgValMetAsnPheGlu 160
DB 424 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGACGCGCTGTGAAGGCGAG 483
QY 161 IleHisblybAlaLeuLybLeuLybAspGlyGlyVHlserTyrlleuValGluPheLybSerIle 180
DB 484 ATCCAAAGGCGCTGAAGCTGAAGAGAGCGGCGCACTACCTGTGGAGTTCAAGTTCATC 543
QY 181 TyrlleAlaLyblybProValGlnLeuProGlyTyrglybValIleAspSerlyblybAsp 200
DB 544 TACATGGCCAAAGAGCGCGCTGACATGCGGCTCACTACGATGAGCTCCAAAGCTGAC 603
QY 201 IleThrSerHlbaenglybAspTyrglybIleValGluGlnTyrglybArgThrGluGlyArg 220
```

```
DB 604 ATCACTCCCAACAGAGACTACATCTGTGAGAGTACGAGCGACCGAGGCGCGC 663
QY 221 HisHlsleuPheLeu 225
DB 664 CACCACTGTTCCTG 678
```

## RESULT 24

US-10-121-258-23  
Sequence 23, Application US/10121258  
Publication No. US20030059835A1  
GENERAL INFORMATION:

```
/ APPLICANT: Tsien, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ FILE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: UC083.LCP2CPI
/ CURRENT APPLICATION NUMBER: US/10/121,258
/ CURRENT FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 23
/ LENGTH: 681
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: nucleotide sequence encoding Dered with mammalian
/ OTHER INFORMATION: codon usage
US-10-121-258-23
```

## Alignment Scores:

```
Prod. No.: 2,2e-142 Length: 681
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: Gaps: 0
```

US-10-006-922a-12 (1-225) x US-10-121-258-23 (1-681)

```
QY 1 MetArgSerSerlybAnValIlelybGluPheMetArgPheValArgMetGluGly 20
DB 4 GTCGGCTCTCCCAAGAACGTCAACAGAGTTTCATGGCTTCAAGGCGCATGAGGCG 63
QY 21 ThrValAsnGlyVHlsgLupheGluIleGluGlyGluGlyArgProTyrgLugly 40
DB 64 ACCGTAAAGCCCAAGATCGAGATCGAGGCGAGGCGCGCCCTTACGAGGCG 123
QY 41 HisAsnThrValIleuLeuValThrlybGlyGlyProLeuProPheAlaTPAspIle 60
DB 124 CACAAACCGGTGAAGCTGAAGTGAACCAAGGCGGCCCCCTGCCCCCTTGGGACATC 183
QY 61 LeuSerProGlnPheGlnTyrglySerlybValTyrglybHlspProAlaAspIlePro 80
DB 184 CTGTCCCCCAAGTTCAGATCGAGCTCCAGATGATGAAGACCCCGCGACATCCCC 243
QY 81 AspTyrllyblyLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100
DB 244 GACTACAAAGAGCTGTCTCCCGAGGCGCTTCAAGTGGAGCGCGCTGATGAACCTTCGAG 303
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerleuGlnAspGlyCysPheIleTy 120
DB 304 GACGGCGCGCTGTGATGACCTGAACCAAGACTCTCCCTGCAAGGCGGCTCTTCATTTAC 363
QY 121 LysValIlePheIleGlyValIleAsnPheProSerAspGlyProValMetGlnlyblybThr 140
DB 364 AAGGTGAAGTTCACTGGCGCTGAACCTTCCCTCGAAGCGGCCCGTAAGCAAGAAACACC 423
QY 141 MetGlyTrpGluIleAsnSerPheProGluArgLeuTyrglybTrpGluArgValMetAsnPheGlu 160
```

Db 424 ATGGCTGGAGGACCTCCACAGCGCCCTGTACCCCGGAGCGCGTGTGAAGGGCGAG 483  
|||  
Qy 161 IIEHISLYSAIALEULYSELYSAPRGLYGTHIAITRYLEUVALGIUHPHELYSSERTILE 180  
|||  
Db 484 ATCCCAAGAGCCCTGAAGCTGAAGAGCGGCGCACTACCTGTGTGAGTTCAAGTCCATC 543  
|||  
Qy 181 TYRMEIAIALYSELYSAPROVALGINLEUPROGLYTRYTRYTRYTRYVALASPSERYLSLEUAAP 200  
|||  
Db 544 TACATGGCCCAAGAGCCCTGTGACGCTGCCGCTACTACTACTGTGACTCCAAAGCTGAGAC 603  
|||  
Qy 201 IIEHRSERHISASNGLUASPTRYTHRIILEVALGIUGINTRYGLUARGTHRGUJLYARG 220  
|||  
Db 604 ATCACTCCCAACACGAGGACTACACATCGTGAAGAGTAAGAGGCGACGAGGGCCGCG 663  
|||  
Qy 221 HISHISLEUPHELEU 225  
|||  
Db 664 CACCACTGTCTCTG 678  
|||

## RESULT 25

US-10-311-030-8  
; Sequence 8, Application US/10311030  
; Publication No. US20040171107A1  
; GENERAL INFORMATION:  
; APPLICANT: Nelson, David  
; APPLICANT: Zamiatra, Elize  
; APPLICANT: Tsien, Roger  
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS  
; FILE REFERENCE: 15916-032US1  
; CURRENT APPLICATION NUMBER: US/10/311,030  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: PCT/US01/04625  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: US 60/184,732  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(678)  
US-10-311-030-8

## Alignment Scores:

Pred. No.: 2.2e-142 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-311-030-8 (1-681)

Qy 1 MetAIGSERISLYSAENVALIILEYSGIUPHEMELARGPHELYSVALARGMETGLUGLY 20  
|||  
Db 4 GTGAGAGAGCAGCAAGACGTGATCAAGAGGTTCATGAGGTTCAGGTGGCATGAGGGCC 63  
|||  
Qy 21 THIRVALAENGILHISGLUPHEGLIUEGLUGIUGIUGIUGIUGIUGIUGIUGIUGIUGI 40  
|||  
Db 64 ACCGTGAGAGCCGACGAGTTCAGATCGAGGCGGAGGCGGAGGCGGAGGCGCTTACAGGGCC 123  
|||  
Qy 41 HISAENITHVALIYSELYSAPROVALGINLEUPROGLYTRYTRYTRYTRYVALASPSERYLSLEUAAP 60  
|||  
Db 124 CACAACACCGTGAAGTTAAGTGACCAAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 183  
|||  
Qy 61 LEUSERPROGLINPHEGLINTRYGLYSELYSVALIYTRYVALYSHISPROLAAAPLLEPO 80  
|||  
Db 184 CTGAGCCCCCAAGTTCAGATCGGAGCAAGGTGTGATCGTGMAACACCCCGCGACATCCCC 243  
|||

Qy 81 APTRYLYSLYSELYSAPROGLIUGIYPHELYSTRIPGLUARGVALMETASNPHEGLI 100  
|||  
Db 244 GACTCAAGAGAGCTGAGCTTCCCGAGGGCTTCAGTGGAGAGGGGTGATGAACTTCAG 303  
|||  
Qy 101 APTGLYGLYVALIATHIRVALITHRGINAPSERISERYLSLEUGIINAPGLIYASPHIELLEYR 120  
|||  
Db 304 GACGGCGCGGTGTGACCGTGACCCAGGACAGAGCCCTGACAGAGCGGCTGCTTCACTTAC 363  
|||  
Qy 121 LYSVALIYSPHEILEGLYVALIASPHEPROSERAPRGLYPROVALMETGLIULYSELYSTR 140  
|||  
Db 364 AAGGTGAAGTTATCTGCGGTGAACCTCCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 423  
|||  
Qy 141 METGLITRGLUAIASERTHRIUGARGLEUTRYPROAAGAPGLYVALIYSELYSGLIUGI 160  
|||  
Db 424 ATGGCTGGAGAGCTTCCACGAGGCGCTGTATACCCCGGACGCGGCTGTGAAGGGCGAG 483  
|||  
Qy 161 IIEHISLYSAIALEULYSELYSAPRGLYGTHIAITRYLEUVALGIUHPHELYSSERTILE 180  
|||  
Db 484 ATCCCAAGAGCCCTGAAGCTGAAGAGCGGCGGCGCACTACCTGTGTGAGTTCAAGTCCATC 543  
|||  
Qy 181 TYRMEIAIALYSELYSAPROVALGINLEUPROGLYTRYTRYTRYTRYVALASPSERYLSLEUAAP 200  
|||  
Db 544 TACATGGCCCAAGAGCCCTGTGACGCTGCCGCTACTACTACTGTGACTCCAAAGCTGAGAC 603  
|||  
Qy 201 IIEHRSERHISASNGLUASPTRYTHRIILEVALGIUGINTRYGLUARGTHRGUJLYARG 220  
|||  
Db 604 ATCACTCCCAACACGAGGACTACACATCGTGAAGAGTAAGAGGCGACGAGGGCCGCG 663  
|||  
Qy 221 HISHISLEUPHELEU 225  
|||  
Db 664 CACCACTGTCTCTG 678  
|||

## RESULT 26

US-10-931-304-3  
; Sequence 3, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Campbell, Robert  
; APPLICANT: Tsien, Roger  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
US-10-931-304-3

Alignment Scores:  
Pred. No.: 2.2e-142 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-3 (1-681)

```

;
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding Dsred with mammalian
; OTHER INFORMATION: codon usage
US-10-931-304-23

Alignment Scores:
Pred. No.: 2,2e-142      Length: 681
Score: 1210.00          Matches: 224
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.56%      Mismatches: 0
Query Match: 99.67%      Indels: 0
DB: 9                      Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-23 (1-681)

QY 1 MetArgSerSerLybAnValIleLygJuphMeArpPheLybValArgMetGluGly 20
   :
   :
   :
Db 4 GTGGCTCTCTCCAGAGACGTCATCAAGAGTTCAATCGCTTCAAGGTGGCATGAGGGC 63
QY 21 ThrValAsnGlyYHIsGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40
   :
   :
   :
Db 64 ACCGTGACGGCCACAGTTCAGATCGAGGGGAGGGCCGCGCCCTACGAGGGC 123
QY 41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60
   :
   :
   :
Db 124 CACAAACACGTTGAAGCTGAAGGTGACCAAGGGGGGGCCCTTGGCTGGACATC 183
QY 61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80
   :
   :
   :
Db 184 CTGTCCCCCAGTTCCAGTACGGCTCCAAAGTGTACGTAAAGCAACCCCGCATCTCCC 243
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100
   :
   :
   :
Db 244 GACTACAAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGAATCTCGAG 303
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
   :
   :
   :
Db 304 GACGGGGCGCGTGGTGAACCGTGAACCAAGACTCTCTCCCTGCAAGACGGCTCTTCACTAC 363
QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybThr 140
   :
   :
   :
Db 364 AAGGTGAAGTTCATCGCGCGTGAATCTTCCCTCCGACGGCCCGTAAATGCAAGAAAGACC 423
QY 141 MetGlyTTrpGluAspSerThrGluArgLeuTyrProArgAspGlyValLeuLybGlyGlu 160
   :
   :
   :
Db 424 ATGGGCTGGAGGGCTTCCACCGAGCGCTGTACCCCGCACGGCGGTGAAGGGGGAG 483
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180
   :
   :
   :
Db 484 ATCCACAAGGCCCTTGAAGCTGAAGAGACGGGCGCACCACTCTGTGGAGTTCAAGTTCATC 543
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200
   :
   :
   :
Db 544 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAGCTCAAGCTGCAC 603
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
   :
   :
   :
Db 604 ATCACTCCCAACAAGAGACTACACCATGTGTGAGAGTACGAGCCGACGAGGGCGCG 663
QY 221 HisHisLeuPheLeu 225
   :
   :
   :
Db 664 CACCACTGTCTCTG 678

RESULT 27
US-10-931-304-23
; Sequence 23, Application US/10931304
; Publication No. US2050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23

```

```

;
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding Dsred with mammalian
; OTHER INFORMATION: codon usage
US-10-931-304-23

Alignment Scores:
Pred. No.: 2,2e-142      Length: 681
Score: 1210.00          Matches: 224
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.56%      Mismatches: 0
Query Match: 99.67%      Indels: 0
DB: 9                      Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-23 (1-681)

QY 1 MetArgSerSerLybAnValIleLygJuphMeArpPheLybValArgMetGluGly 20
   :
   :
   :
Db 4 GTGGCTCTCTCCAGAGACGTCATCAAGAGTTCAATCGCTTCAAGGTGGCATGAGGGC 63
QY 21 ThrValAsnGlyYHIsGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40
   :
   :
   :
Db 64 ACCGTGACGGCCACAGTTCAGATCGAGGGGAGGGCCGCGCCCTACGAGGGC 123
QY 41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60
   :
   :
   :
Db 124 CACAAACACGTTGAAGCTGAAGGTGACCAAGGGGGGGCCCTTGGCTGGACATC 183
QY 61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80
   :
   :
   :
Db 184 CTGTCCCCCAGTTCCAGTACGGCTCCAAAGTGTACGTAAAGCAACCCCGCATCTCCC 243
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100
   :
   :
   :
Db 244 GACTACAAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGAATCTCGAG 303
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
   :
   :
   :
Db 304 GACGGGGCGCGTGGTGAACCGTGAACCAAGACTCTCTCCCTGCAAGACGGCTCTTCACTAC 363
QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybThr 140
   :
   :
   :
Db 364 AAGGTGAAGTTCATCGCGCGTGAATCTTCCCTCGACGGCCCGTAAATGCAAGAAAGACC 423
QY 141 MetGlyTTrpGluAspSerThrGluArgLeuTyrProArgAspGlyValLeuLybGlyGlu 160
   :
   :
   :
Db 424 ATGGGCTGGAGGGCTTCCACCGAGCGCTGTACCCCGCACGGCGGTGAAGGGCGAG 483
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180
   :
   :
   :
Db 484 ATCCACAAGGCCCTTGAAGCTGAAGAGACGGGCGCACCACTCTGTGGAGTTCAAGTTCATC 543
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200
   :
   :
   :
Db 544 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAGCTCAAGCTGCAC 603
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
   :
   :
   :
Db 604 ATCACTCCCAACAAGAGACTACACCATGTGTGAGAGTACGAGCCGACGAGGGCGCG 663
QY 221 HisHisLeuPheLeu 225
   :
   :
   :
Db 664 CACCACTGTCTCTG 678

RESULT 28
US-10-311-030-11
; Sequence 11, Application US/10311030
; Publication No. US20040171107A1
; GENERAL INFORMATION:
; APPLICANT: Nelson, David
; APPLICANT: Zamtata, Elize

```

[illegible]

Db	622	ATCACGACCCACACGACGACTTACCACTCGTGGAGCACTACGAGACCGAGGCGAGG	681
Qy	221	HisHisLeupheLeu 225	
Db	682	CACGACCTGTTCTTG 696	
RESULT 29			
US-10-311-030-12/c			
; Sequence 12, Application US/10311030			
; Publication No. US20040171107A1			
; GENERAL INFORMATION:			
; APPLICANT: Nelson, David			
; APPLICANT: Zamata, Elize			
; APPLICANT: Tsien, Roger			
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS			
; FILE REFERENCE: 15916-032US1			
; CURRENT APPLICATION NUMBER: US/10/311, 030			
; PRIOR FILING DATE: 2002-12-10			
; PRIOR APPLICATION NUMBER: PCT/US01/04625			
; PRIOR FILING DATE: 2001-02-13			
; PRIOR APPLICATION NUMBER: US 60/184,732			
; PRIOR FILING DATE: 2000-02-23			
; NUMBER OF SEQ ID NOS: 12			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 12			
; LENGTH: 713			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: synthetic construct			
US-10-311-030-12			
Alignment Scores:			
Pred. No.: 2,34e-142 Length: 713			
Score: 1210.00 Matches: 224			
Percent Similarity: 100.00% Conservative: 1			
Best Local Similarity: 99.56% Mismatches: 0			
Query Match: 99.67% Indels: 0			
DB: 7 Gaps: 0			
US-10-006-922A-12 (1-225) x US-10-311-030-12 (1-713)			
Qy	1	MeCaRgSerSerLybAenValIlleLySgluPheMeArGpHeLyValArgMeCgluGly 20	
Db	692	GTGAGGAGCAACAAGACGTATCAAGAGATTCAAGAGTTCAAGGTGCGATGAGGAGC 633	
Qy	21	ThrValaenGlyHisGluPheGluIlleGluGlyGluGlyArgProGlyArgGly 40	
Db	632	ACCGTGAACGGCCACGAATTGAGATTCAGAGGCGGAGGCGGAGGCGGACGCTTACGAGGCG 573	
Qy	41	HisAenThrValLybLeuLybValThrLySglYgLyProLeuProPheAlaTrpAapIle 60	
Db	572	CACAACACCGTGAAGCTTAAAGTGAACCAAGGGCGGCCCCCTTCCTCGCGGACATC 513	
Qy	61	LeuSerProGluPheGluInTrGlySerLySValYrValLybHisProAlaAapIlePro 80	
Db	512	CTGAGCCCCCAGTTCACATGACGAGCAAGAGGTGTAAGACACCCCGCGACATCCCC 453	
Qy	81	AspTrIlybLybLeuSerPheProGluGlyPheLySTpGluArgValMetAenPheGlu 100	
Db	452	GACTACAAAGAAAGCTGAGCTTCCCGAGGGCTTCAAGTGGGAAGAGGTGATGAATTCGAG 393	
Qy	101	AspGlyGlyValValThrValThrGluAenPheSerSerLeuGluAapGlyCySPheIleYr 120	
Db	392	GACGGCGGCGGTGTGACCGTGAACCAAGACACAGACCTGCGAGGACGGCTTCACTTAC 333	
Qy	121	LySValLySPheIleGlyValAaSPheProSerAapGlyProValMetGluInLybYrThr 140	
Db	332	AAAGTGAAGTTCATCGGCGGTGAATCTTCCCGACGCGAGCGGCCGTATGCAAGAAAGACC 273	
Qy	141	MetGlyTrpGluAaSerThrGluArgLeuYrProArgAapGlyValLeuLySglYgLy 160	

Db 272 ATGGGCTGGAGGCTCCACGAGCGCTGTACCCCGGACGCGGTGTGAAGGCGAG 213  
QY 116H1SLYSLALeuleuLeuYbApG1yG1YH1sTYrLeuValGluPhelysSer11e 180  
Db 212 ATCCACAGGCGCTTGAAGTGAAGACGCGCGCCACTACCTGTGAGTTCAAGTCCATC 153  
QY 161 TTMetAlaLYbLeuProValGlnLeuProG1YTYrTYrTYrValAspSerLYbLeuAap 200  
Db 152 TACATGGCCAGAGAGCCCGTGCAGCTGCCGCGCTACTACGTGACTCCAAAGCTGGAC 93  
QY 201 I1eHrSerH1AspGluAspTYrThrI1eValGluGlnTYrGluAspThrGluG1YAYG 220  
Db 92 ATCACCAGCCACAGACGAGACTACACATGTGAGACAGTACAGAGACCGAGGCGAG 33  
QY 221 H1SH1SLeupheLeu 225  
Db 32 CACCACTGTTCTCTG 18  
RESULT 30  
US-10-152-296-1  
/ Sequence 1, Application US/10152296  
/ Publication No. US20030077730A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Peelle, Beau  
/ APPLICANT: Rigel Pharmaceuticals, Incorporated  
/ TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
/ FILE REFERENCE: 021044-000110US  
/ CURRENT APPLICATION NUMBER: US/10/152,296  
/ CURRENT FILING DATE: 2002-12-10  
/ PRIOR APPLICATION NUMBER: US 60/291,871  
/ PRIOR FILING DATE: 2001-05-18  
/ NUMBER OF SEQ ID NOS: 2  
/ SOFTWARE: PatentIn Ver. 2.1  
/ SEQ ID NO 1  
/ LENGTH: 723  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Description of Artificial Sequence:mammalian  
/ OTHER INFORMATION: codon-optimized variant (DERED) of Discosoma sp.  
/ OTHER INFORMATION: "red" red fluorescent protein (RFP)  
/ NAME/KEY: CDS  
/ LOCATION: (1)..(723)  
/ OTHER INFORMATION: DERRD  
US-10-152-296-1  
Alignment Scores:  
Pred. No.: 2,39e-142 Length: 723  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 5 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-152-296-1 (1-723)  
QY 1 MetArgSerSerLYbAspValI1eLYbGluPhemeRxpPhelysValArgMetGluG1Y 20  
Db 4 GTGCGCTCTCCAGAGACGTATCAGAGAGTTCATGCGCTTCAAGTGGCATGGAAGGC 63  
QY 21 ThValAsnG1YH1eGluPhelGluI1eGluG1Yg1uG1Yg1uG1YArProTYrG1uG1Y 40  
Db 64 ACCGTGAACGCCACGAGTTGAGATCGAGGGCGAGGGCGCGCCCTACGAGGCG 123  
QY 41 H1AspThrValLYbLeuLYbValThLYbG1Yg1YProLeuProPheAlaTPAspI1e 60  
Db 124 CACAAACACCGTGAAGCTGAAGTGAACCAAGGCGCGCCCTCGCCCTCGCTGGACATC 183  
QY 61 LeuSerProGlnPhelG1NYrG1YSerLYbValTYrValLYbSH1sProAlaAspI1ePro 80  
Db 184 CTGTCCCCCAATTCCAGTACGGCTCCAGGTGTACGTGAAGCACCCCGCGACATCCCC 243

QY 81 AspTYrLYbLYbLeuSerPheProG1uG1YPhelysTYrPG1uArgValMetAspNheGlu 100  
Db 244 GACTACAGAGAGACTCTCTTCCCGAGGCTTCAAGTGGAGCGCGCTATGAACTTCGAG 303  
QY 101 AspG1Yg1YValIValThrValThrGluAspSerSerLeuGlnAspG1YCyAspHe11eTYr 120  
Db 304 GACGCGGCGGTGTGATCCGTGACCCAGAGACTCTCTCTGACAGACGCGTGTTCATCTAC 363  
QY 121 LYbValLYbPheI1eG1YValAspNheProSerAspG1YProValMetGluLYbLYbThr 140  
Db 364 AAGGTGAAGTTCATCATGGGTGAACTTCCCTCCAGCGCGCCGTATGATGAGAGAGACC 423  
QY 141 MetG1YTPG1uAlaSerThrG1uArgLeuTYrProArgAspG1YValLeuLYbG1Yg1u 160  
Db 424 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGGACGCGCTGTGAAGGCGAG 483  
QY 161 I1eH1SLYSLALeuleuLeuYbApG1yG1YH1sTYrLeuValGluPhelysSer11e 180  
Db 484 ATCCACAGGCGCTTGAAGTGAAGACGCGCGCCACTACCTGTGAGTTCAAGATATC 543  
QY 181 TTMetAlaLYbLeuProValGlnLeuProG1YTYrTYrTYrValAspSerLYbLeuAap 200  
Db 544 TACATGGCCAGAGAGCCCGTGCAGCTGCCGCGCTACTACTACGTGACTCCAAAGCTGGAC 603  
QY 201 I1eHrSerH1AspGluAspTYrThrI1eValGluGlnTYrGluAspThrGluG1YAYG 220  
Db 604 ATCACCCTCCACAGACGAGACTACACATGTGAGACAGTACAGAGCGCACCGAGGCGCG 663  
QY 221 H1SH1SLeupheLeu 225  
Db 664 CACCACTGTTCTCTG 678  
RESULT 31  
US-10-739-656-1  
/ Sequence 1, Application US/10739656  
/ Publication No. US20040126850A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Peelle, Beau  
/ APPLICANT: Rigel Pharmaceuticals, Incorporated  
/ TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
/ FILE REFERENCE: 021044-000110US  
/ CURRENT APPLICATION NUMBER: US/10/739,656  
/ CURRENT FILING DATE: 2003-12-17  
/ PRIOR APPLICATION NUMBER: US/10/152,296  
/ PRIOR FILING DATE: 2002-12-10  
/ PRIOR APPLICATION NUMBER: US 60/291,871  
/ PRIOR FILING DATE: 2001-05-18  
/ NUMBER OF SEQ ID NOS: 2  
/ SOFTWARE: PatentIn Ver. 2.1  
/ SEQ ID NO 1  
/ LENGTH: 723  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Description of Artificial Sequence:mammalian  
/ OTHER INFORMATION: codon-optimized variant (DERED) of Discosoma sp.  
/ OTHER INFORMATION: "red" red fluorescent protein (RFP)  
/ NAME/KEY: CDS  
/ LOCATION: (1)..(723)  
/ OTHER INFORMATION: DERRD  
US-10-739-656-1  
Alignment Scores:  
Pred. No.: 2,39e-142 Length: 723  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 7 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-739-656-1 (1-723)

```
QY      1 MetArgSerSerLyAsnValIleYsgIuphMeCArgPheLyValArgMetGluGly 20
      4 GTGGCTCTCTCCAAAGACTCATCAAGAGTTCAATGCGCTTCAAGGTGGCATGAGGGC 63
DB      21 ThrValaengIYHieGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
      64 ACCGTGAACGGCCAGAGTTGAGATCGAGGGCGAGGGCGGGCCGCCCTTACAGAGGGC 123
QY      41 HIsaenThrValLyLeuLyValThrLyseGlyYProLeuProPheLaITraPheIle 60
      124 CACAACACCGTAAAGCTGAAGGTGACCAAGGGCGGGCCCTTCCCTTCCCTGGAGATC 183
DB      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHIsaProLaAspIlePro 80
      184 CTGTCTCCCACTTCCAGTACGGCTCCAAAGGTGATGATGAACACCCGCCCACTCCCC 243
QY      81 AepTyrLyLyLeuSerPheProGluGlyPheLySTPGLuArgValMeCAsnPheGlu 100
      244 GACTACAAAGAGCTGTCTTCCCGAGGGGCTTCAAGTGGAGCGCGTGAATTCGAG 303
DB      101 AepGlyGlyValValThrValThrGlnAepSerSerLeuGlnAepGlyCysePheIleTyr 120
      304 GACGGCGGGGTGTGACCGGTGACCAAGACTCTCTCCGAGAGAGCGGTCTTCAATAC 363
QY      121 LySValLyPheIleGlyValAsnPheProSerAepGlyProValMetGlnLySlyThr 140
      364 AAGGTGAAGTTCAATGGCGCTGAATCTCCCTCCGACGGCCCGGTAAATGAGAAGAACCC 423
DB      141 MetGlyTTPGluIleSerThrGluArgLeuTyrProArgAepGlyValLeuLySlyGlu 160
      424 ATGGCTGGAGAGCTCCCAACGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAG 483
QY      161 IleHIslyValaLeuLyLeuLyPheArgGlyYHIsTyrLeuValGlnPheLySerIle 180
      484 ATCCCAAGAGCCCTTAAGGTGAAGAGCGGGCCACTACTGTGTGAGTTCAAGAGTATC 543
DB      181 TyrMetAlaLyLyPProValGlnLeuProGlyTyrTyrTyrValAspSerLySleuAep 200
      544 TPCATGGCCAAAGAACCCGTGCAAGCTGCCCGCTACTACTAGTGGACTCCAAAGCTGAGC 603
QY      201 IleThrSerHIsaengIuAepTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
      604 ATCACCTCCCAAGAGAGACTACACATCTGTGAGAGTACGAGGCGACCGAGGGCCGC 663
DB      221 HIsHIsleuPheLeu 225
      664 CACCACTGTTCCTG 678
QY
DB
RESULT 32
US-10-214-932-51
; Sequence 51, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhwan
; APPLICANT: LEE, Yong Jik
; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
; FILE REFERENCE: AP802/US
; CURRENT APPLICATION NUMBER: US/10/214,932
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
; NAME/KEY: gene
; LOCATION: (1)..(1638)
; OTHER INFORMATION: Gene for RFP-PS(N1a protease):AtOEP7:GFP
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (1)..(1635)
; OTHER INFORMATION: RFP:PS(N1a protease):AtOEP7:GFP hybrid protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: CDS: RFP(1..678); CDS: N1a protease proteolytic site(700..711); C
; OTHER INFORMATION: DS: AtOEP7(712..900); CDS: GFP(922..1635)
US-10-214-932-51

Alignment Scores:
Pred. No.: 7,46e-142 Length: 1638
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
Gaps: 0

US-10-006-922A-12 (1-225) x US-10-214-932-51 (1-1638)
QY      1 MetArgSerSerLyAsnValIleYsgIuphMeCArgPheLyValArgMetGluGly 20
      4 GTGGCTCTCTCCAAAGACTCATCAAGAGTTCAATGCGCTTCAAGGTGGCATGAGGGC 63
DB      21 ThrValaengIYHieGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
      64 ACCGTGAACGGCCAGAGTTGAGATCGAGGGCGAGGGCGGGCCGCCCTTACAGAGGGC 123
QY      41 HIsaenThrValLyLeuLyValThrLyseGlyYProLeuProPheLaITraPheIle 60
      124 CACAACACCGTAAAGCTGAAGGTGACCAAGGGCGGGCCCTTCCCTTCCCTGGAGATC 183
DB      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHIsaProLaAspIlePro 80
      184 CTGTCTCCCACTTCCAGTACGGCTCCAAAGGTGATGATGAACACCCGCCCACTCCCC 243
QY      81 AepTyrLyLyLeuSerPheProGluGlyPheLySTPGLuArgValMeCAsnPheGlu 100
      244 GACTACAAAGAGCTGTCTTCCCGAGGGGCTTCAAGTGGAGCGCGTGAATTCGAG 303
DB      101 AepGlyGlyValValThrValThrGlnAepSerSerLeuGlnAepGlyCysePheIleTyr 120
      304 GACGGCGGGGTGTGACCGGTGACCAAGACTCTCTCCGAGAGAGCGGTCTTCAATAC 363
QY      121 LySValLyPheIleGlyValAsnPheProSerAepGlyProValMetGlnLySlyThr 140
      364 AAGGTGAAGTTCAATGGCGCTGAATCTCCCTCCGACGGCCCGGTAAATGAGAAGAACCC 423
DB      141 MetGlyTTPGluIleSerThrGluArgLeuTyrProArgAepGlyValLeuLySlyGlu 160
      424 ATGGCTGGAGAGCTCCCAACGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAG 483
QY      161 IleHIslyValaLeuLyLeuLyPheArgGlyYHIsTyrLeuValGlnPheLySerIle 180
      484 ATCCCAAGAGCCCTTAAGGTGAAGAGCGGGCCACTACTGTGTGAGTTCAAGAGTATC 543
DB      181 TyrMetAlaLyLyPProValGlnLeuProGlyTyrTyrTyrValAspSerLySleuAep 200
      544 TPCATGGCCAAAGAACCCGTGCAAGCTGCCCGCTACTACTAGTGGACTCCAAAGCTGAGC 603
QY      201 IleThrSerHIsaengIuAepTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
      604 ATCACCTCCCAAGAGAGACTACACATCTGTGAGAGTACGAGGCGACCGAGGGCCGC 663
DB      221 HIsHIsleuPheLeu 225
      664 CACCACTGTTCCTG 678
QY
DB
RESULT 33
US-10-214-932-75
; Sequence 75, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhwan
```

```
/ APPLICANT: KIM, Dae Heon
/ APPLICANT: LEE, Yong Jik
/ TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
/ FILE REFERENCE: APB02/US
/ CURRENT APPLICATION NUMBER: US/10/214,932
/ CURRENT FILING DATE: 2002-08-08
/ NUMBER OF SEQ ID NOS: 133
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 75
/ LENGTH: 1647
/ TYPE: DNA
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Synthetic Sequence
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (1)..(1647)
/ OTHER INFORMATION: Gene for RFP:PS(HIV-1 protease):AtoEP7:GFP hybrid protein
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1644)
/ OTHER INFORMATION: RFP:PS(HIV-1 protease):AtoEP7:GFP hybrid protein
/ NAME/KEY: misc-feature
/ OTHER INFORMATION: CDS: RFP(1..678); CDS: HIV-1 protease proteolytic site(700..720);
/ OTHER INFORMATION: CDS: AtoEP7(721..909); CDS: GFP(931..1644)
US-10-214-932-75

Alignment Scores:
Pred. No.: 7,52e-142 Length: 1647
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-214-932-75 (1-1647)
QY 1 MetArgSerSerIysAsnValIleIleGluPheMetArgPheIysValArgMetGluGly 20
DB 4 GTGGCGCTCTCCCAAGAACGTGATCAAGAGTTTCATGGCTTCAAGGTGCGCATGAGAGGC 63
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB 64 ACCGTGAACGGCCACGAGTTCAGAGTCAGAGGCGAGGGCCGCCCTCAAGAGGCG 123
QY 41 HisAsnThrValIysLeuIysValThrIleGlyGlyProLeuProPheIleAspIle 60
DB 124 CACAAACCGGTGAAGCTGAAGTGAACCAAGGCGGCGCCCTGCTGCGCTGGAGCATC 183
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProIleAspIlePro 80
DB 184 CTGTCCCCCGACGTTCCAGTACGCTCCAGGTGACGTGAAGACACCCCGCGACATCCC 243
QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTyrGluArgValMetAsnPheGlu 100
DB 244 GACTACAAAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGGATGAATCTCGAG 303
QY 101 AspGlyGlyValIleThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB 304 GACGCGCGCGGTGACCGTGAACCAAGACCTCCCTCGAGGACGGCTGCTTATCATAC 363
QY 121 LysValIysPheIleGlyValIleAsnPheProSerAspGlyProValMetGlnIysLeuThr 140
DB 364 AAGGTGAAGTTCACTCGGCTGAATCTCCCTCCAGCGCCCCCGTAACTCAAGAAAGACC 423
QY 141 MetGlyTyrGluIleAspSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyIle 160
DB 424 ATGGGCTGGAGGCTCCACCGAGGCTGTACCCCGCGAGCGGCTGTAAAGGCGAG 483
QY 161 IleHisIysValIleLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180
DB 484 ATCCAAAGGCGCTGAAGCTGAAGGACGGCGGCACTACCTGTGGAGTTCAAGTCCATC 543
```

```
QY 181 TyrMetAlaIysIysProValGlnLeuProGluTyrTyrTyrValAspSerIysLeuAsp 200
DB 544 TACATGGCCAAAGAGCTGTGAGCGGCTACTACTAGTGAATCCAGCTGAGAC 603
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluIysArg 220
DB 604 ATCACTCCCAACAGAGGACTACACCATGTGGAGCACTAGAGGCAACGAGGCGCCG 663
QY 221 HisHisLeuPheLeu 225
DB 664 CACCACTGTCTCTG 678

RESULT 34
US-10-161-403-29
/ Sequence 29, Application US/10161403
/ Publication No. US20030119104A1
/ GENERAL INFORMATION:
/ APPLICANT: Perkins, Edward
/ APPLICANT: Perez, Carl
/ APPLICANT: Lindenbaum, Michael
/ APPLICANT: Greene, Amy
/ APPLICANT: Leung, Josephine
/ APPLICANT: Fleming, Elena
/ APPLICANT: Stewart, Sandra
/ APPLICANT: Shellard, Joan
/ TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
/ FILE REFERENCE: 24601-420
/ CURRENT APPLICATION NUMBER: US/10/161,403
/ CURRENT FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: 60/294,758
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: 60/366,891
/ PRIOR FILING DATE: 2002-03-21
/ NUMBER OF SEQ ID NOS: 129
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 29
/ LENGTH: 4692
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: pds:red1-N1 plasmid from Clontech
US-10-161-403-29

Alignment Scores:
Pred. No.: 3,23e-141 Length: 4692
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-161-403-29 (1-4692)
QY 1 MetArgSerSerIysAsnValIleIleGluPheMetArgPheIysValArgMetGluGly 20
DB 682 GTGGCGCTCTCCCAAGAACGTGATCAAGAGTTTCATGGCTTCAAGGTGCGCATGAGAGGC 741
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB 742 ACCGTGAACGGCCACGAGTTCAGAGTCAGAGGCGAGGGCGCGCCCTCAAGAGGCG 801
QY 41 HisAsnThrValIysLeuIysValThrIleGlyGlyProLeuProPheIleAspIlePro 60
DB 802 CACAAACCGGTGAAGCTGAAGTGAACCAAGGCGGCGCCCTGCTGCGCTGGAGATC 861
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProIleAspIlePro 80
DB 862 CTGTCCCCCGACGTTCCAGTACGCTCCAGGTGACGTGAAGACACCCCGCGACATCCC 921
QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTyrGluArgValMetAsnPheGlu 100
DB 922 GACTACAAAGAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTGAATGAATCTCGAG 981
```

Qy	101	AspGlyGlyValValThrValThrglnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	982	GACGGCGCGTGTATCAGCTGACCCAGAGACTCTCTCTCAGAGACGGCTGCTTCACTTAC	1041
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	1042	AAGGTGAAGTTGATCGGGCTGAACCTTCCCTCGAGCGGCCGGTATGCAAGAAAGAAC	1101
Qy	141	MetGlyTTPGlnAlaSerThrGlnArgLeuTyrProAspGlyValLeuLysGlyGlu	160
Db	1102	ATGGGCTGTGGAGGCTCCACCCAGCGCCTGTACCCCGGAGCGGCTGTGTGAAGGGCGAG	1161
Qy	161	IleHisLysValAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGlnPheLysSerIle	180
Db	1162	ATCCCAAGGCCCTTAAGCTGAAGAGCGCGGCGACACTGCTGGTGAAGTTCAAGTCCATC	1221
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrValAspSerLysLeuAsp	200
Db	1222	TACATGGCCAGAAACCCGTGAGCTGCCCGGCTACTTACGTGACCTCCAAAGCTGGAC	1281
Qy	201	IleThrSerHisAsnGlnAspTyrThrIleValGlnGlnTyrGlnArgThrGlnGlyArg	220
Db	1282	ATCACCCTCCCAACAGAGACTACACCACTGTGGAGCACTAACAGAGCGCACCGAGGCGCCG	1341
Qy	221	HisHisLeuPheLeu	225
Db	1342	CACCAACCTGTTCTTG	1356

```

RESULT 35
US-10-433-640-16
; Sequence 16, Application US/10433640
; Publication No. US20040115792A1
; GENERAL INFORMATION:
; APPLICANT: Lichtenberg-Frate, Heila
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF
; TITLE OF INVENTION: COMPLEX ENVIRONMENTAL CONTAMINATIONS
; FILE REFERENCE: 1487/3
; CURRENT APPLICATION NUMBER: US/10/433,640
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/14610
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: DE 10061872.3
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 4692
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: seqecne of vector pdsRed1-N1
US-10-433-640-16

Alignment Scores:
Pred. No.: 3,23e-141 Length: 4692
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-433-640-16 (1-4692)

```

QY 1 MetAaSeSeSerIyAaAaVallIeIyAgiUpHeMeAaGPeIyAaValAaMeGluGly 20  
Db 682 GTGGCTCTCTCCAGAAAGTCATCAAGAGATTCAATGCTTCAGAGTCCGATGGAAGGC 7414

QY 21 ThrValAaMgIyHiSeIupHeGluIeGluIyGluIyAaGProTyGluGly 40  
Db 742 ACGGTAAACGGCCACAGATTCCAGATCGAGGCGAGGCGAGGCGCGCTCTACGAAAGGC 8014

QY 41 HlaaThrValIyAaLeuIyValThrIyAgiGlyProIeProPhaIaTrpAspIle 60

Db	802	CACAAKACCGTGAAGCTGAAGGTACCAAGGCGGCCCCCTCGCCCTTCGCTTGGGACATC	861
QY	61	LeuSerProGlnPheGlnTYrGlySerLysValTYrValLYsHISProAlaSp1LePro	80
Db	862	CTGTCCCCCCAGTTCCAGTACGCGCTCCAGGTGTACGTGAAGCACCCCGCCGACATCCCC	921
QY	81	AspTYrLysLysLeuSerPheProGluGlyPheLysTrpGluAryValMetAsnPheGlu	100
Db	922	GACTACAAAGAAAGCTGTCTCTCCCGAAGGCTTCAGATGGAGCGCGTCAATAACTTCAG	981
QY	101	AspGlyLysValThrValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheLeTYr	120
Db	982	GACGGCGGCGGTGGACCGTGACCAGACACTCTCCCTGCAGACACGGCTGCTTCATCTAC	1041
QY	121	LysValLysPheLeileGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	1042	AAGGTGAAGTTCACCGCGGTGAACCTTCCTCCACGCGCCCGTAAATGCAGAAAGAAC	1101
QY	141	MetGlyTrpGluAlaSerThrGluAryLeuTYrProAryAspGlyValLeuLysGlyGlu	160
Db	1102	ATGGGCTGGAGGCGCTCACACCGAGCGCTGTACCCCGCCACACGCGCTGTAAAGGCGAG	1161
QY	161	LeHisLysAlaLeuLysLysLysAspGlyGlyHisTYrLeuValGluPheLysSerLe	180
Db	1162	ATCCACAAGGCGCTGAAGCTGAAGACGCGGCGCACACTACCTGTGTGAGTTCAAGTCCATC	1221
QY	181	TYrMetSerLysLysProValGlnLeuProGlyTYrTYrTYrValAspSerLysLeuAsp	200
Db	1222	TACATGGCCAAAGAGCCCGTGACACTCCCGGCTCACTACGTGAGATCTCCAGAGTGAAC	1281
QY	201	LeIleHisSerHisAsnGluAspTYrThrLeValGluGlnTYrGluAryThrGluGlyAry	220
Db	1282	ATCACTCCCCCAACAGACGAGACTACACCATCGTGGAGCATACGAGCGCACCGAGGGCGCG	1341
QY	221	HisHisLeuPheLeu 225	
Db	1342	CACCACTGTTCCTG 1356	

```

RESULT 36
US-11-006-076-29
Sequence 29, Application US/11006076
Publication No. US20050181506a1
GENERAL INFORMATION:
APPLICANT: Perkins, Edward
APPLICANT: Perez, Carl
APPLICANT: Lindenbaum, Michael
APPLICANT: Greene, Amy
APPLICANT: Leung, Josephine
APPLICANT: Fleming, Elena
APPLICANT: Stewart, Sandra
APPLICANT: Shellard, Joan
TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
FILE REFERENCE: 17084-022002/420B
CURRENT APPLICATION NUMBER: US/11/006,076
CURRENT FILING DATE: 2004-12-06
PRIOR APPLICATION NUMBER: 60/294,758
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 60/366,891
PRIOR FILING DATE: 2002-03-21
NUMBER OF SEQ ID NOS: 129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29
LENGTH: 4692
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pdsred1-N1 plasmid from Clontech
US-11-006-076-29

```

Alignment scores:		
Pred. No.:	3.23e-141	Length: 4692
Score:	1210.00	Matches: 224



Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922a-12 (1-225) x US-11-006-076-29 (1-4692)

```
QY 1 MetArgSerSerLyAsnValIleLySGIuphemeTarGpHeLyValArgMetGluGly 20
DB : : : : :
QY 682 GAGCGCTCTCCAAAGAACGTCAACAGAGATTCAATCGCTTCAAGTGCAGTAGAGGAC 741
DB 21 ThrValaenGlyHIGluPhelGluIleGluGluGluGluGluGluGluGluGluGluGlu 40
QY 742 ACCGTGAAAGCGGCAAGAGTTCAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 801
DB 41 HIsenThValLyLeuLyValThLySGLyGlyProLeuProPheLaTrPaPile 60
QY 802 CACMACACCGTGAAGCTTGAAGTGAACMACAGGCGGCGCGCTTCCCTGCGGACATC 861
DB 61 LeuSerProGlnPheGlnTyrgLySerLyValTyTrValLyHIsProLaAapIlePro 80
QY 862 CTGTCCCTCCAAAGTTCAGATCGGCTCCAAAGTGTACGTGAAGACCCCGCATCCCC 921
DB 81 AspTyTrLyLeuLeuSerPheProGluGlyPheLyTrGluArgValMetAsnPhelGly 100
QY 922 GACTACAAAGAGTGTCTTCTCCCGAGGCGCTTCAAGTGGAGCGCGTGAATGAACCTTCAG 981
DB 101 AaPGIyGlyValValaThrValThrgInaAspSerSeLeuGlnaAspGlyCySpheIleTy 120
QY 982 GACGGGCGGAGTGTGACCGTGAACCGAGACTCTCTCTGAGAGCGGCTGCTTCACTAC 1041
DB 121 LyValaLySpheIleGlyValaAsnPhProSerAspGlyProValMetGlnLySlyThr 140
QY 1042 AAGGTGAAGTTCATCGGCGGTGAACCTTCCCTCCGACGGCGCGCTTAATGCAAGAAAGACC 1101
DB 141 MetGlyTrpGluAlaSerThrgIuArgLeuTyTrProArgAspGlyValaLeuLySGLyGlu 160
QY 1102 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGCAGCGCGCTGAAAGGCGAG 1161
DB 161 HIsLyValaLeuLySleuLyAspGlyGlyHIsTyTrLeuValaGluPhelysSerIle 180
QY 1162 ATCCAAAGGCTTGAAGCTGAAGAGAGCGGCGCACTACTGAGATTCAAGTCCATC 1221
DB 181 TyMetAlaLySlySProValGlnLeuProGlyTyTrTyTrValaAspSerLySleuAsp 200
QY 1222 TACATGGCCAAAGAGCGCGTGAAGTCCCGGCTACTACTACGTGACCTCAAGCTGAGAC 1281
DB 201 IleThSerHIsaenGluAspTyTrThrIleValaGluGlnTyTrGluArgThrgIuGlyArg 220
QY 1282 ATCACTCCCAACAAGAGACTACACCATGTGTGAGAGTACGAGCGCACCGAGGCGGC 1341
DB 221 HIsHIsleuPhelGlu 225
DB 1342 CACCACTGTCTCTG 1356
```

RESULT 37  
US-10-169-050-46

Sequence 46, Application US/10169050  
Publication No. US20050071891A1  
GENERAL INFORMATION:  
APPLICANT: THRESHER, RON  
APPLICANT: HINDS, LYNN  
APPLICANT: HARDY, CHRIS  
APPLICANT: WHYARD, STEVE  
APPLICANT: VIGNABAJAN, SOMA  
APPLICANT: GREWE, PETER MARTIN  
APPLICANT: PATIL, JAMAHAR  
TITLE OF INVENTION: REPRESSIBLE STERILITY OF ANIMALS  
FILE REFERENCE: 4050, 001500  
CURRENT APPLICATION NUMBER: US/10/169, 050  
PRIOR APPLICATION NUMBER: PCT/AU00/0  
PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: P04884  
PRIOR FILING DATE: 1999-12-24  
NUMBER OF SEQ ID NOS: 63  
SOFTWARE: Patent version 3.2  
SEQ ID NO 46  
LENGTH: 5436  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: pSPFM6  
US-10-169-050-46

#### Alignment Scores:

Pred. No.:	3,976-141
Score:	1210.00
Percent Similarity:	100.00%
Best Local Similarity:	99.56%
Query Match:	99.67%
DB:	9
Length:	5436
Matches:	224
Conservative:	1
Mismatches:	0
Indels:	0
Gaps:	0

US-10-006-922a-12 (1-225) x US-10-169-050-46 (1-5436)

```
QY 1 MetArgSerSerLyAsnValIleLySGIuphemeTarGpHeLyValArgMetGluGly 20
DB : : : : :
QY 1426 GAGCGCTCTCCAAAGAACGTCAACAGAGATTCAATCGCTTCAAGTGCAGTAGAGGAC 1485
DB 21 ThrValaenGlyHIGluPhelGluIleGluGluGluGluGluGluGluGluGluGluGlu 40
QY 1486 ACCGTGAAAGCGGCAAGAGTTCAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1545
DB 41 HIsenThValLyLeuLyValThLySGLyGlyProLeuProPheLaTrPaPile 60
QY 1546 CACMACACCGTGAAGCTTGAAGTGAACMACAGGCGGCGCGCTTCCCTGCGGACATC 1605
DB 61 LeuSerProGlnPheGlnTyrgLySerLyValTyTrValLyHIsProLaAapIlePro 80
QY 1606 CTGTCCCTCCAAAGTTCAGATCGGCTCCAAAGTGTACGTGAAGACCCCGCATCCCC 1665
DB 81 AspTyTrLyLeuLeuSerPheProGluGlyPheLyTrGluArgValMetAsnPhelGly 100
QY 1666 GACTACAAAGAGTGTCTTCTCCCGAGGCGCTTCAAGTGGAGCGCGTGAATGAACCTTCAG 1725
DB 101 AaPGIyGlyValValaThrValThrgInaAspSerSeLeuGlnaAspGlyCySpheIleTy 120
QY 1726 GACGGGCGGAGTGTGACCGTGAACCGAGACTCTCTCTGAGAGCGGCTGCTTCACTAC 1785
DB 121 LyValaLySpheIleGlyValaAsnPhProSerAspGlyProValMetGlnLySlyThr 140
QY 1786 AAGGTGAAGTTCATCGGCGGTGAACCTTCCCTCCGACGGCGCGCTTAATGCAAGAAAGACC 1845
DB 141 MetGlyTrpGluAlaSerThrgIuArgLeuTyTrProArgAspGlyValaLeuLySGLyGlu 160
QY 1846 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGCAGCGCGCTGAAAGGCGAG 1905
DB 161 HIsLyValaLeuLySleuLyAspGlyGlyHIsTyTrLeuValaGluPhelysSerIle 180
QY 1906 ATCCAAAGGCTTGAAGCTGAAGAGAGCGGCGCACTACTGAGATTCAAGTCCATC 1965
DB 181 TyMetAlaLySlySProValGlnLeuProGlyTyTrTyTrValaAspSerLySleuAsp 200
QY 1966 TACATGGCCAAAGAGCGCGTGAAGTCCCGGCTACTACTACGTGACCTCAAGCTGAGAC 2025
DB 201 IleThSerHIsaenGluAspTyTrThrIleValaGluGlnTyTrGluArgThrgIuGlyArg 220
QY 2026 ATCACTCCCAACAAGAGACTACACCATGTGTGAGAGTACGAGCGCACCGAGGCGGC 2085
DB 221 HIsHIsleuPhelGlu 225
DB 2086 CACCACTGTCTCTG 2100
```

RESULT 38

US-10-001-189-45  
Sequence 45, Application US/10001189

```
/ Publication No. US20020173634A1
/ GENERAL INFORMATION:
/ APPLICANT: FRASER JR., MALCOLM J.
/ APPLICANT: LI, XU
/ APPLICANT: BEAM, TERESA
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
/ TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
/ FILE REFERENCE: 835910-92098
/ CURRENT APPLICATION NUMBER: US/10/001,189
/ CURRENT FILING DATE: 2001-10-30
/ PRIOR APPLICATION NUMBER: 60/244,984
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: 60/244,677
/ PRIOR FILING DATE: 2000-10-31
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 45
/ LENGTH: 6984
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:
US-10-001-189-45

Alignment Scores:
Pred. No.: 5,62e-141 Length: 6984
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: Gaps: 0

US-10-006-922A-12 (1-225) x US-10-001-189-45 (1-6984)

QY 1 MetArgSerSerLySerAnValIleLySgIuphMetArgPheLySerValArgMetGluGly 20
DB 2974 GTGCGCTCTCCAAAGACTCATCAAGAGCTTCATCGCTTCAAGTGGCATGAGGGC 3033

QY 21 ThrValaEngLyHsGluPhGluIleGluGlyGluGlyValArgProTyrgLugly 40
DB 3034 ACCGGAAACGGCCACGAGTTCAGATCGAGGCGGAGGCGCGCCCTACGAGGCGC 3093

QY 41 HsAenThrValaLyLeuLySerValaThryLySgLyProLeuProPheLaATrpaSple 60
DB 3094 CACAACACCGTAAGCTGAAGGTGACCAAGGGGCGCCCTCGCTTGGCGGACATC 3153

QY 61 LeuSerProGlnPheGlnTyrgLySerLySgValaTyrgValaLyHsIaProLaAspIlePro 80
DB 3154 CTGTCCCTCCCAATTCAGTAAGGCTCCAGGCTGTAAGTGAAGACCCCGCAATCCCC 3213

QY 81 AspTyrgLyLeuLySerPheProGluGlyPheLySTrPGluArgValaMetAenPheGlu 100
DB 3214 GACTCAAGAAAGCTGTCTTCCCGGAGGCTTCAGTGGAGGCGCTGATGAATTCGAG 3273

QY 101 AaPGLyGlyValaValaThryValaThrglnAaPserSerLeuGlnAaSPGlyCyarPheIleTy 120
DB 3274 GACGGGGGGGTGGTGAACCGTGAACCGGACTCTCTCCAGGACCGGCTGCTTCATTCAC 3333

QY 121 LySValaLyPheHleGlyValaAenPheProSerAaSPGlyProValaMetGlnLySlyTh 140
DB 3334 AAGGTAATTCATCTCGGCTGAATCTTCCCTCCGACGCGCCCGTAATGCAAGAAAGACC 3393

QY 141 MetGlyTrPGluAaSerThrglnArgLeuTyrgProArgAaSPGlyValaLeuLySgLyGlu 160
DB 3394 ATGGGCTGGGAGGCGCTCCACCGAGCGCTGTACCCCGGACGCGCGCTGCTGAAGGGCGAG 3453

QY 161 IleHsLyValaLeuLySerLySgAaSPGlyGlyHsIaTyrgLeuValaGlnPheLySserIle 180
DB 3454 ATCCCAAGAGGCGCTTGAAGCTGAAGAGCGGCGGCACTACCTGATGAGTTCAAGTCCATC 3513

QY 181 TyrgMetAlaLyLeuAaProValaGlnLeuProGluTyrgTyrgValaAaPserLySleuAaP 200
```

```
DB 3514 TACATGGCCAAAGAGCCCGTAGCTGCCGCGCTACTACTAGTGAATCCCAAGCTGGAC 3573

QY 201 IleThrSerHsAenGluAaSPGlyThryIleValaGluGlnTyrgLuarGThrgLuarGlyArg 220
DB 3574 ATCACTCCCTCCCAAGAAAGAACTAACATCTGTGAGAGTGAAGCGGACCGAGGCGCGC 3633

QY 221 HsHsIaLeuPheLeu 225
DB 3634 CACCACTGTCTCTG 3648

RESULT 39
US-10-169-050-20/c
/ Sequence 20, Application US/10169050
/ Publication No. US20050071891A1
/ GENERAL INFORMATION:
/ APPLICANT: THRESHER, RON
/ APPLICANT: HINDS, LYNN
/ APPLICANT: HARDY, CHRIS
/ APPLICANT: WHYARD, STEVE
/ APPLICANT: VIGNARAJAN, SOMA
/ APPLICANT: GREWE, PETER MARTIN
/ APPLICANT: PATIL, JAMNABHAR
/ TITLE OF INVENTION: REPRESSIBLE STERILITY OF ANIMALS
/ FILE REFERENCE: 4050.001500
/ CURRENT APPLICATION NUMBER: US/10/169,050
/ CURRENT FILING DATE: 2002-06-24
/ PRIOR APPLICATION NUMBER: PCT/AU00/0
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: P04884
/ PRIOR FILING DATE: 1999-12-24
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 20
/ LENGTH: 7910
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: pBMT (dHSP) - RFP-OHoxDS/BH
US-10-169-050-20

Alignment Scores:
Pred. No.: 6.69e-141 Length: 7910
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: Gaps: 0

US-10-006-922A-12 (1-225) x US-10-169-050-20 (1-7910)

QY 1 MetArgSerSerLySerAnValIleLySgIuphMetArgPheLySerValaArgMetGluGly 20
DB 2705 GTGCGCTCTCCAAAGACTCATCAAGAGTTCATGCGCTTCAAGTGGCATGAGGGC 2646

QY 21 ThrValaEngLyHsGluPhGluIleGluGlyGluGlyValArgProTyrgLugly 40
DB 2645 ACCGTGAACGGCCACGAGTTCAGATCGAGGCGGAGGCGCGCCCTACGAGGGC 2586

QY 41 HsAenThrValaLyLeuLySerValaThryLySgLyProLeuProPheLaATrpaSple 60
DB 2585 CACAACACCGTGAAGCTGAAGGTGACCAAGGGGCGCCCTGCTGCGCTGGGACATC 2526

QY 61 LeuSerProGlnPheGlnTyrgLySerLySgValaTyrgValaLyHsIaProLaAspIlePro 80
DB 2525 CTGTCCCTCCCAATTCAGTAAGGCTCCAGGCTGTAAGTGAAGACCCCGCAATCCCC 2466

QY 81 AspTyrgLyLeuLySerPheProGluGlyPheLySTrPGluArgValaMetAenPheGlu 100
DB 2465 GACTCAAGAAAGCTGTCTTCCCGGAGGCTTCAGATGGAACCGGCTGATGAATTCGAG 2406

QY 101 AaPGLyGlyValaValaThryValaThrglnAaPserSerLeuGlnAaSPGlyCyarPheIleTy 120
```

```

Db      2405 GACGCGCGCGGTGGAGACCGGTGAGACCCAGAGACTCTCCCTCGAGAGACGGCTGCTCATCTAC 2346
Qy      121  LyValAllyPheIleGlyValAlaenPheProSerAspGlyProValMetGlnLysLeuThr 140
Db      2345 AAGGTGAAGTTTCATCGCGCGTGAACCTTCCCTCCGACGGCCCCGTAATGCGAAGAGACGC 2286
Qy      141  MetGlyTrpGluAlaSerThrGluArgLeuLysProArgAspGlyValIleLysGlyVal 160
Db      2285 ATGGCGTGGAGGGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGCTGTAAAGGGCGAG 2266
Qy      161  IleHisValAlaLeuLysLeuLysAspGlyGlyHisTrpLeuValGluPheLysSerIle 180
Db      2225 ATCCACAAAGCCCTCGAAGCTGTAAGAGCGGGCGGCACATCACTGCGTGAAGTTCAAGTTCATC 2166
Qy      181  TyrMetAlaLysValysProValGlnLeuProGlyTrpTrpValAspSerLysLeuAsp 200
Db      2165 TACATGCGCAAGAGACCGCGTCACTGCGCCGCGTACTCTACGAGACCTCCAAAGCTGAC 2106
Qy      201  IleThrSerHisAsnGluAspTrpThrIleValGluGlnTrpArgThrGluGlyArg 220
Db      2105 ATCACTCCCTCCCAACGAGAGACTTACCACTGCGAGCAGTACGAGCGCACGAGGCGGCGCCG 2046
Qy      221  HisHisLeuPheLeu 225
Db      2045 CACCACTGTCTCTG 2031

RESULT 40
; Sequence 20, Application US/10471065
; Publication No. US20040197855A1
GENERAL INFORMATION:
APPLICANT: Prof. Dr. Wiem Mueller, Lisa
TITLE OF INVENTION: Test system for the determination of genotoxicities
FILE REFERENCE: P59532
CURRENT APPLICATION NUMBER: US/10/471,065
CURRENT FILING DATE: 2003-09-04
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 9320
TYPE: DNA
ORGANISM: Artificial sequence
FEATURES:
OTHER INFORMATION: Description of the artificial sequence: Plasmid
OTHER INFORMATION: p5-Puro-CMV-(N'-EGFP)-CMV-Red-(EGFP-EJ)
FEATURES:
NAME/KEY: misc.feature
LOCATION: (1)..(1592)
OTHER INFORMATION: Retroviral vector p5NM
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3378)..(3392)
OTHER INFORMATION: Retroviral vector p5NM
FEATURE:
NAME/KEY: misc.feature
LOCATION: (5527)..(9320)
OTHER INFORMATION: Retroviral vector p5NM
FEATURE:
NAME/KEY: gene
LOCATION: (1617)..(2216)
OTHER INFORMATION: Purocmv resistance gene from pRetron (Clontech,
OTHER INFORMATION: Palo Alto, CA, USA)
FEATURE:
NAME/KEY: promoter
LOCATION: (2267)..(2848)
OTHER INFORMATION: CMV promoter from pEGFP-N1 (Clontech, Palo Alto,
OTHER INFORMATION: CA, USA)
FEATURE:
NAME/KEY: gene
LOCATION: (2906)..(3348)
OTHER INFORMATION: N'-EGFP, derived from EGFP from pEGFP-N1
OTHER INFORMATION: (Clontech, Palo Alto, CA, USA)
FEATURE:

```

[illegible]

```
RESULT 41
US-10-609-019-4
/ Sequence 4, Application US/10609019
/ Publication No. US20040197910A1
/ GENERAL INFORMATION:
/ APPLICANT: Cooper, Richard K.
/ APPLICANT: Cadd, Gary G.
/ APPLICANT: Fioretti, William C.
/ APPLICANT: DeBeer, Kenneth F.
/ TITLE OF INVENTION: Gene Regulation in Transgenic Animals Using a Transposon-Based
/ FILE REFERENCE: 51687-0101 (51687-287015)
/ CURRENT APPLICATION NUMBER: US/10/609,019
/ CURRENT FILING DATE: 2003-06-26
/ PRIOR APPLICATION NUMBER: US 60/392,415
/ PRIOR FILING DATE: 2002-06-26
/ PRIOR APPLICATION NUMBER: US 60/441,392
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,377
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,502
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,405
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,447
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,381
/ PRIOR FILING DATE: 2003-01-21
/ SOFTWARE: Patent in version 3.2
/ SEQ ID NO 4
/ LENGTH: 9658
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic
US-10-609-019-4

Alignment Scores:
Pred. No.:      8,83e-141      Length:      9658
Score:          1210.00        Matches:      224
Percent Similarity: 100.00%    Conservative: 1
Best Local Similarity: 99.56%  Mismatches:    0
Query Match:    99.67%        Indels:       0
DB:             8             Gaps:         0

US-10-006-922A-12 (1-225) x US-10-609-019-4 (1-9658)
QY      1 MetArgSerSerLyAsnValIlleYsgIuphMetArgPheLyValArgMetGluGly 20
DB      5099 GTGGCTCTCCAAAGACTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGAC 5158
QY      21 ThrValAsnGlyHIsGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40
DB      5159 ACCGGAAAGGCGACAGATTGAGATCGAGGCGAGGCGGCGGCCCTTACAGAGGCG 5218
QY      41 HisAsnThrValIleLeuLyValThrLyGlyGlyProLeuProPheAlaTrpAspIle 60
DB      5219 CACAACACCGGTGAAGCTGAAGTGAACCAAGGCGGCGCCCTCGCTCGGAGCATC 5278
QY      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValIleHisProAlaAspIlePro 80
DB      5279 CTGTCCCCCAAGTTCCAGTACGGCTCCAGGTGTACGTGAAGCAACCCCGACATCCCC 5338
QY      81 AspTyrLyValLeuSerPheProGluGlyPheLySTirGluArgValMetAsnPheGlu 100
DB      5339 GACTACAAAGACTGTCTTCCCGAGGCGCTTCAAGTGGAGCGCGTATGAACTTCGAG 5398
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      5399 GACGGCGGGGTGTGACCGTGAACCCAGGACTCTCTCCCTGACAGACGCGCTTCAATC 5458
QY      121 LySValLyPheIlleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140
```

```
DB      5459 AAGGTGAAGTTCATGGGGGTGAACCTTCCCTCCAGCGGCCCGGTAAATGACAGAAAGACC 5518
QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProAlaAspGlyValIleuLyGlyIu 160
DB      5519 ATGGCTGGAGGCGCTTCCACCGAGCGCTGTACCCCGCGAGCGGTGCTGAAGGCGAG 5578
QY      161 IleHisLyValIleuLyLeuLyAsnGlyGlyIleTyrIleuValGluPheLySerIle 180
DB      5579 ATCCCAAGGCGCTGTAACTGAAGACCGCGGCACTACCTGTGTGAGATTCAAGTTCATC 5638
QY      181 TyrMetAlaLySlyPheProValGlnLeuProGluTyrTyrTyrValAspSerLyAsnAsp 200
DB      5639 TACATGGCCAAAGAGCCCGTGTGAGCTGCCGCGCTACTACTAGTGACTCCAAAGCTGAGC 5698
QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB      5699 ATCACTCCCAACACGAAGACTACCATCTGTGAGCGATGAGACCGACCGAGGCGCG 5758
QY      221 HisHisLeuPheLeu 225
DB      5759 CACCACTGTTCCTG 5773

RESULT 42
US-10-609-019-3
/ Sequence 3, Application US/10609019
/ Publication No. US20040197910A1
/ GENERAL INFORMATION:
/ APPLICANT: Cooper, Richard K.
/ APPLICANT: Cadd, Gary G.
/ APPLICANT: Fioretti, William C.
/ APPLICANT: DeBeer, Kenneth F.
/ TITLE OF INVENTION: Gene Regulation in Transgenic Animals Using a Transposon-Based
/ FILE REFERENCE: 51687-0101 (51687-287015)
/ CURRENT APPLICATION NUMBER: US/10/609,019
/ CURRENT FILING DATE: 2003-06-26
/ PRIOR APPLICATION NUMBER: US 60/392,415
/ PRIOR FILING DATE: 2002-06-26
/ PRIOR APPLICATION NUMBER: US 60/441,392
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,377
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,502
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,405
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,447
/ PRIOR FILING DATE: 2003-01-21
/ PRIOR APPLICATION NUMBER: US 60/441,381
/ PRIOR FILING DATE: 2003-01-21
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: Patent in version 3.2
/ SEQ ID NO 3
/ LENGTH: 9678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic
US-10-609-019-3

Alignment Scores:
Pred. No.:      8,85e-141      Length:      9678
Score:          1210.00        Matches:      224
Percent Similarity: 100.00%    Conservative: 1
Best Local Similarity: 99.56%  Mismatches:    0
Query Match:    99.67%        Indels:       0
DB:             8             Gaps:         0

US-10-006-922A-12 (1-225) x US-10-609-019-3 (1-9678)
QY      1 MetArgSerSerLyAsnValIlleYsgIuphMetArgPheLyValArgMetGluGly 20
DB      5099 GTGGCTCTCCAAAGACTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGAC 5158
QY      21 ThrValAsnGlyHIsGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40
DB      5159 ACCGGAAAGGCGACAGATTGAGATCGAGGCGAGGCGGCGGCCCTTACAGAGGCG 5218
QY      41 HisAsnThrValIleLeuLyValThrLyGlyGlyProLeuProPheAlaTrpAspIle 60
DB      5219 CACAACACCGGTGAAGCTGAAGTGAACCAAGGCGGCGCCCTCGCTCGGAGCATC 5278
QY      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValIleHisProAlaAspIlePro 80
DB      5279 CTGTCCCCCAAGTTCCAGTACGGCTCCAGGTGTACGTGAAGCAACCCCGACATCCCC 5338
QY      81 AspTyrLyValLeuSerPheProGluGlyPheLySTirGluArgValMetAsnPheGlu 100
DB      5339 GACTACAAAGACTGTCTTCCCGAGGCGCTTCAAGTGGAGCGCGTATGAACTTCGAG 5398
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      5399 GACGGCGGGGTGTGACCGTGAACCCAGGACTCTCTCCCTGACAGACGCGCTTCAATC 5458
QY      121 LySValLyPheIlleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140
```

```

Db      5119  GTGCGCTCTCCAGAAAGTCATCAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 5178
      21  ThrValaenGIYHlAgIuPheGluIleGIuGIYGIuGIYArgProTYrGIuGIY 40
      5179  ACCGTGAAGGCGCAAGATTCCAGATCGAGGGGAGGGGCGGCCCTTACGAGGGC 5238
      41  HlAenThValIyLeuYleuYValThrIyGlyGIYProLeuProPheAlaTPAapIle 60
      5239  CACAAACCGTGAAGCTGAAGGTGACCAAGGCGGGCCCTTCCGCTGGGAGCATC 5298
      61  LeuSerProGlnPheGlnTYrGIYSerIyValTYrValIyHlAProIlaAapIlePro 80
      5299  CTGTCCCCCAGTTCCAGTACGAGCTCCAGAGTGTAGTGAAGACCCCGCGACATCCCC 5358
      81  AApTYrLyLeuYleuSerPheProGIuGIYPhelYsTPGIuArgValMetAenPheGlu 100
      5359  GACTACAAAGAGCTGTCTTCTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAATCTCGAG 5418
      101  AApGIYGIYValIyThrValThrGlnAAsPseSerIeuGlnAapGIYCySpheIleTYr 120
      5419  GACGGGCGGTGTGACCGTGAACCGACCAAGACTCTCTCCGTGAGAGCGGCTGCTTCACTAC 5478
      121  LysValIyPheHlIegIYValAenPheProSerAapGIYProValMetGlnIySlySThr 140
      5479  AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCGAGCGGCCGTAAATGCAAGAAAGACC 5538
      141  MetGIYTPGIuIaAseThrGIuArgLeuTYrProArgAapGIYValIeulysGIYGIu 160
      5539  ATGGGCTGGAGGCTCTCCACCGAGCGCTTACCCCGACGCGCGTGAAGGGCGAG 5598
      161  HlAenThValIyLeuYleuYsaPseGIYGIYHlAeTYrLeuValGIuPheIySerIle 180
      5599  ATCCACAAAGGCGCTGAAGCTGAAGAGAGCGGCGCACTACCTGGTGAAGTTCAAGTCATC 5658
      181  TyrMetAlaIySlyPseProValGlnLeuProGIYTYrTYrTYrValAAsPseIySleuAap 200
      5659  TACATGGCCAAAGAGCGCGGTGACGTGCCCGGCTACTACTACGTGAGCTCAAGCTGAGC 5718
      201  HlAeThSerHlAenGIuAapTYrThrIleValGIuGlnTYrGIuArgThrGIuGIYArg 220
      5719  ATCACTCCCAACAAGAGACTACATCATGTGAGAGCATGACGCCACGAGGGCGCG 5778
      221  HlAeHlAeupheIeu 225
      5779  CACCACTGTCTCG 5793

```

RESULT 43  
 US-10-609-019-2  
 ; Sequence 2, Application US/10609019  
 ; Publication No. US20040197910A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cooper, Richard K.  
 ; APPLICANT: Cadd, Gary G.  
 ; APPLICANT: Fioretti, William C.  
 ; APPLICANT: DeBoer, Kenneth F.  
 ; TITLE OR INVENTION: Gene Regulation in Transgenic Animals Using a Transposon-Based  
 ; TITLE OR INVENTION: Vector  
 ; FILE REFERENCE: 51687-0101 (51687-287015)  
 ; CURRENT APPLICATION NUMBER: US/10/609, 019  
 ; PRIOR FILING DATE: 2003-06-26  
 ; PRIOR APPLICATION NUMBER: US 60/392, 415  
 ; PRIOR FILING DATE: 2002-06-26  
 ; PRIOR APPLICATION NUMBER: US 60/441, 392  
 ; PRIOR FILING DATE: 2003-01-21  
 ; PRIOR APPLICATION NUMBER: US 60/441, 377  
 ; PRIOR FILING DATE: 2003-01-21  
 ; PRIOR APPLICATION NUMBER: US 60/441, 502  
 ; PRIOR FILING DATE: 2003-01-21  
 ; PRIOR APPLICATION NUMBER: US 60/441, 405  
 ; PRIOR FILING DATE: 2003-01-21  
 ; PRIOR APPLICATION NUMBER: US 60/441, 447  
 ; PRIOR FILING DATE: 2003-01-21  
 ; PRIOR APPLICATION NUMBER: US 60/441, 381

```

      ; PRIOR FILING DATE: 2003-01-21
      ; NUMBER OF SEQ ID NOS: 43
      ; SOFTWARE: Patentin version 3.2
      ; SEQ ID NO 2
      ; LENGTH: 10263
      ; TYPE: DNA
      ; ORGANISM: Artificial Sequence
      ; FEATURE:
      ; OTHER INFORMATION: Synthetic
      US-10-609-019-2

Alignment Scores:
Pred. No.: 9 61e-141
Score: 1210.00
Percent Similarity: 100.00%
Best Local Similarity: 99.56%
Query Match: 99.67%
DB: 8 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-609-019-2 (1-10263)

QY      1  MetArgSerSerLySaenValIleIyGIuPheMetArgPheIySaValArgMetGIuGIY 20
      :::
      5725  GTGCGCTCTCCAAAGAACGTATCAAGAGATTCAATGCGCTTCAAGGTGCGCATGAGAGGC 5784
      21  ThrValaenGIYHlAgIuPheGluIleGIuGIYGIuGIYArgProTYrGIuGIY 40
      5785  ACCGTGAAGGCGCAAGTTCAGATCGAGGAGGAGGGGCGGCCCTTACGAGGGC 5844
      41  HlAenThValIyLeuYleuYsaPseGIYGIYProLeuProPheAlaTPAapIle 60
      5845  CACAAACCGTGAAGCTGAAGGTGACCAAGGCGGGCCCTTCCGCTGGAGCATC 5904
      61  LeuSerProGlnPheGlnTYrGIYSerIyValTYrValIyHlAProIlaAapIlePro 80
      5905  CTGTCCCCCAGTTCCAGTACGAGCTCCAAAGTGTAGTGAAGACCCCGCGACATCCCC 5964
      81  AApTYrLyLeuYleuSerPheProGIuGIYPhelYsTPGIuArgValMetAenPheGlu 100
      5965  GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGGCGGTGATGAATCTTCGAG 6024
      101  AApGIYGIYValIyThrValThrGlnAAsPseSerIeuGlnAapGIYCySpheIleTYr 120
      6025  GACGGGCGGTGTGACCGTGAACCGAGACTCTCTCCGTGAGAGCGGCTCTTCACTTAC 6084
      121  LysValIyPheHlIegIYValAenPheProSerAapGIYProValMetGlnIySlySThr 140
      6085  AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCGAGCGGCCGTAAATGCAAGAAAGACC 6144
      141  MetGIYTPGIuIaAseThrGIuArgLeuTYrProArgAapGIYValIeulysGIYGIu 160
      6145  ATGGGCTGGAGGCTCTCCACCGAGCGCTGTATCCCCCGACAGCGGCTGTAAGGGCGAG 6204
      161  HlAenThValIyLeuYleuYsaPseGIYGIYHlAeTYrLeuValGIuPheIySerIle 180
      6205  ATCCACAAAGGCGCTGAAGCTGAAGAGAGGCGCACTACTGTGTGAGTTCAAGTCCATC 6264
      181  TyrMetAlaIySlyPseProValGlnLeuProGIYTYrTYrTYrValAAsPseIySleuAap 200
      6265  TACATGGCCAAAGAGCGCGGTGACGTGCCCGGCTACTACTACGTGAGCTCAAGCTGAGC 6324
      201  HlAeThSerHlAenGIuAapTYrThrIleValGIuGlnTYrGIuArgThrGIuGIYArg 220
      6325  ATCACTCCCAACAAGAGACTACATCATGTGAGAGCATGACGCCACCGAGGGCGCG 6384
      221  HlAeHlAeupheIeu 225
      6385  CACCACTGTCTCG 6399

```

RESULT 44  
 US-10-315-920-3  
 ; Sequence 3, Application US/10315920  
 ; Publication No. US20030175809A1

```

/ GENERAL INFORMATION:
/ APPLICANT: Pradekov, Arcady Fedorovich
/ APPLICANT: Teredikh, Alexey
/ TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
/ TITLE OF INVENTION: FOR THEIR USE
/ FILE REFERENCE: CLON-077CIP
/ CURRENT APPLICATION NUMBER: US/10/315,920
/ PRIOR FILING DATE: 2002-12-09
/ PRIOR APPLICATION NUMBER: 60/211,607
/ PRIOR FILING DATE: 2000-06-14
/ PRIOR APPLICATION NUMBER: PCT/US01/19097
/ PRIOR FILING DATE: 2001-06-13
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-3

```

```

Alignment Scores:
Pred. No.: 5,22e-142 Length: 678
Score: 1207.00 Matches: 223
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.11% Mismatches: 1
Query Match: 99.42% Indels: 0
DB: 6 Gaps: 0

```

US-10-006-922a-12 (1-225) x US-10-315-920-3 (1-678)

```

QY 1 MetArgSerSerLyAsnValIleLySgluphMetArgPheLyValArgMetGluGly 20
DB 1 ATGGCGTCTCTCAAGAAAGCTGATCAAGAGTTGATCGCTTCAAGTGGCGCATGAGGGC 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGluGluGluGluGluGluGluGluGluGlu 40
DB 61 ACCGTGAACGGCCACAGATTGAGATCGAGGCGAGGGGAGGGCGCCCTTACGAGGGC 120
QY 41 HisAsnThrValIleLeuLyValIleLySglYProLeuProPheAlaTrpAspIle 60
DB 121 CACAACACCGTGAAGCTGAAAGGTGACCAAGGGGGGCGCCCTGCTCCCTGGGACATC 180
QY 61 LeuSerProGlnPheGlnTrpGlySerLySValIYrValIleHisProAlaAspIlePro 80
DB 181 CTGTCCCTCCAGTTCAGTACGAGCTCAAGGTGTACGTAAAGACCCCGCCGACATCCCC 240
QY 81 AspTyrLyLeuLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
DB 241 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAATTCGAG 300
QY 101 AspGlyGlyValValIleThrValIleAspSerSerLeuGlnAspGlyCySPheIleTyr 120
DB 301 GACGGGGGGGTGGCGGACCGGTGACCCGAGACTCTCTCCAGGACCGGCTGCTCATCTAC 360
QY 121 LysValIlePheIleGlyValIleAsnPheProSerAspGlyProValMetGlnLysIleThr 140
DB 361 AAGGTGAAGTTCATCGGCGTGAATCTTCCCTCGAGGGCCCGGTGATGCAAGAAAGACC 420
QY 141 MetGlyTyrGluIleAsnSerThrGluArgLeuTyrProArgAspGlyValLeuIleSglGlu 160
DB 421 ATGGGCTGGGAGGCTCCACCGAGGCGCTGTACCCCGGAGCGGCGTGTGAAGGGCGAG 480
QY 161 IleHisLysLeuAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB 481 ATCCCAAGAGCCCTGAAGCTGAAGGACGGCGGCACTACCTGGTGAAGTTCAAGTCCATC 540
QY 181 TyrMetAlaIleLysLysProValGlnLeuProGluTyrTyrTyrValAspSerLysLeuAsp 200
DB 541 TACATGGCCAAAGAGCCCGGTGACGCTGACCGGCTACTACTAGTGAACCAAGCTGGAC 600
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGluArg 220

```

```

DB 601 ATCACTCCCAAGAAAGAGTACTACACATCGTGAGAGTACGAGCGCACCGAGGGCGGC 660
QY 221 HisIleLeuPheLeu 225
DB 661 CACCACTGTCTCTG 675

```

RESULT 45

US-10-742-828-4  
Sequence 4, Application US/10742828  
Publication No. US20040157294A1

GENERAL INFORMATION:

APPLICANT: Huynh, Gigi

APPLICANT: Heinrich, Gerhard

TITLE OF INVENTION: TRANSGENIC SCREEN AND METHOD FOR SCREENING MODULATORS OF BRAIN-DEF

FILE REFERENCE: US 1353/03 (VA)

CURRENT APPLICATION NUMBER: US/10/742,828

PRIOR FILING DATE: 2003-12-23

PRIOR APPLICATION NUMBER: US 10/306,737

NUMBER OF SEQ ID NOS: 5

SEQ ID NO 4

LENGTH: 7508

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: Unsure

LOCATION:

OTHER INFORMATION: Synthesized

US-10-742-828-4

```

Alignment Scores:
Pred. No.: 4,76e-140 Length: 7508
Score: 1203.00 Matches: 222
Percent Similarity: 99.56% Conservative: 2
Best Local Similarity: 98.67% Mismatches: 1
Query Match: 99.09% Indels: 0
DB: 7 Gaps: 0

```

US-10-006-922a-12 (1-225) x US-10-742-828-4 (1-7508)

```

QY 1 MetArgSerSerLyAsnValIleLySgluphMetArgPheLyValArgMetGluGly 20
DB 2123 GTGGCTCTCTCAAGAAAGCTGATCAAGAGTTGATCGCTTCAAGTGGCGCATGAGGGC 2182
QY 21 ThrValAsnGlyHisGluPheGluIleGluGluGluGluGluGluGluGluGluGluGlu 40
DB 2183 ACCGTGAACGGCCACAGATTGAGATCGAGGCGAGGGGAGGGCGCCCTTACGAGGGC 2242
QY 41 HisAsnThrValIleLeuLyValIleLySglYProLeuProPheAlaTrpAspIle 60
DB 2243 CACAACACCGTGAAGCTGAAAGGTGACCAAGGGGGGCGCCCTGCTCCCTGGGACATC 2302
QY 61 LeuSerProGlnPheGlnTrpGlySerLySValIYrValIleHisProAlaAspIlePro 80
DB 2303 CTGTCCCTCCAGTTCAGTACGAGCTCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 2362
QY 81 AspTyrLyLeuLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
DB 2363 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAATTCGAG 2422
QY 101 AspGlyGlyValValIleThrValIleAspSerSerLeuGlnAspGlyCySPheIleTyr 120
DB 2423 GACGGGGGGGTGGCGGACCGGTGACCCGAGACTCTCTCCAGGACCGGCTGCTCATCTAC 2482
QY 121 LysValIlePheIleGlyValIleAsnPheProSerAspGlyProValMetGlnLysIleThr 140
DB 2483 AAGGTGAAGTTCATCGGCGTGAATCTTCCCTCGAGGGCCCGGTGATGCAAGAAAGACC 2542
QY 141 MetGlyTyrGluIleAsnSerThrGluArgLeuTyrProArgAspGlyValLeuIleSglGlu 160
DB 2543 ATGGGCTGGGAGGCTCCACCGAGGCGCTGTACCCCGGAGCGGCGTGTGAAGGGCGAG 2602

```

```
QY 161 IIEHSLVLALEULYSLEULYSAPG1YGH1STYRLVVALGIUHPHELYSSRT116 180
DB 2603 ATCCACCAAGCCCTGAAAGCTGAAGACGCGGCCCACTACCTGGTGAATTCAAGTCCATC 2662
QY 181 TYRMEALALYLYSPROVALGINLEUPROG1YTYR1TYR1VAL1ASPSELYSLEUAAP 200
DB 2663 TGCATGGCCCAAGAGCCCTGACGTGCGCTGCTACTACTACGTGACACCAAGCTGGAC 2722
QY 201 ILETHSERH1SANG1UAPRYTHR1LEVALG1UG1NTYRG1UARGTHRG1UG1YARG 220
DB 2723 ATCACTCCCAACAAGAGACTACACCATCTGTGACAGTACGAGCGCAACGAGCGCCGC 2782
QY 221 H1SH1SL1EUPHELEU 225
DB 2783 CACCACTGTTCCTG 2797

RESULT 46
US-10-006-922-37
/ Sequence 37, Application US/10006922
/ Publication No. US20020197676A1
/ GENERAL INFORMATION:
/ APPLICANT: Lukyanov, Sergey A.
/ APPLICANT: Fradkov, Arcady F.
/ APPLICANT: Labas, Yulii A.
/ APPLICANT: Matz, Mikhail V.
/ APPLICANT: Terzikikh, Alexey
/ TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and
/ TITLE OF INVENTION: Method for Using the Same
/ FILE REFERENCE: CLON-035CIP
/ CURRENT APPLICATION NUMBER: US/10/006,922
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 09/120,330
/ PRIOR FILING DATE: 1998-12-11
/ PRIOR APPLICATION NUMBER: 09/457,898
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/458,144
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/458,477
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/457,556
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/444,338
/ PRIOR FILING DATE: 1999-11-19
/ NUMBER OF SEQ ID NOS: 46
/ SOFTWARE: ParseSeq for Windows Version 4.0
/ SEQ ID NO 37
/ LENGTH: 681
/ TYPE: DNA
/ ORGANISM: Drosophila species
US-10-006-922-37

Alignment Scores:
Pred. No.: 7,196-141 Length: 681
Score: 1198.00 Matches: 221
Percent Similarity: 99.11% Conservative: 2
Best Local Similarity: 98.22% Mismatches: 2
Query Match: 98.68% Indels: 0
DB: 5 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-006-922-37 (1-681)
QY 1 MetArgSerSerLybAenVal1leLyegLuphMeArpPhelybValArgMetGluG1Y 20
DB 4 GTGGGCTCTCTCCAGAAAGCTCATCAAGAGTTTCAAGGTCCTCAAGGTGCGCATGGAAGGC 63
QY 21 ThrValAng1YH1SG1UPhEG1U1leG1UG1YGLUG1YARGProTYRG1UG1Y 40
DB 64 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGCGAGGCGCGCCCTTACGAGAGGC 123
QY 41 H1SANG1TH1LYSLEULYSLEULYSVAL1Th1RYAG1YGL1YProLeuProPhAla1TriAsp1le 60
DB 124 CACAACACCGTGAAGCTGAAGGTACCAAGAGCGCGCCCTGCTGCTTGGCTGGGACATC 183
```

```
QY 61 LeuSerProGlnPheGln1TYRG1YSer1YSVal1TYRVal1LYSH1SPRO1AlaAsp1lePro 80
DB 184 CTGTCCCTCCCACTTCAAGTACGGCTCCAAAGGTGACGTGAAGCACCCCGCGACATCCCC 243
QY 81 AapTYR1LYSLYLeuSerPheProG1UG1YPhelySTPGLUARGVALMetAnPheGlu 100
DB 244 GACTCAAGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACTTCGAG 303
QY 101 AapG1YGL1YVAL1Th1RYALTh1RG1NAPSerSer1eUG1NAPG1YCYAsPhe1leTYR 120
DB 304 GACGGCGCGGTGCGACCCGTGACCAAGACTCTCTCCGTGACAGACGCGCTTCACTTAC 363
QY 121 LYsVal1LYSPhe1leG1YVal1ANPheProSerAapG1YProVALMetGln1LYSTHR 140
DB 364 AAGGTGAAGTTATGCGGTGAATCTTCCCTCCGACGCGCCGTATGACGAAGAAGAC 423
QY 141 MetG1YTPGLUAlaSerThRG1UARGLeuTYRProARGAPG1YVALLeuYSGL1U 160
DB 424 ATGGGCTGGAGAGCCCTCCACCGAGCGCTGTACCCCGCGACGCGCTGTGAAGCGCGAG 483
QY 161 IIEHSLVLALEULYSLEULYSAPG1YGH1STYRLVVALGIUHPHELYSSRT116 180
DB 484 ACCCAACAAGCCCTGAAAGCTGAAGACGCGCGCACTACTCTGTGAGTTCAAGTCCATC 543
QY 181 TYRMEALALYLYSPROVALGINLEUPROG1YTYR1TYR1VAL1ASPSELYSLEUAAP 200
DB 544 TGCATGGCCCAAGAGCCCTGACGTGCGCTGCTACTACTACTGTAAGCCCAAGCTGGAC 603
QY 201 ILETHSERH1SANG1UAPRYTHR1LEVALG1UG1NTYRG1UARGTHRG1UG1YARG 220
DB 604 ATCACTCCCAACAAGAGACTACACCATCTGTGACAGTACGAGCGCAACGAGCGCCGC 663
QY 221 H1SH1SL1EUPHELEU 225
DB 664 CACCACTGTTCCTG 678

RESULT 47
US-10-332-733-22
/ Sequence 22, Application US/10332733
/ Publication No. US20040106565A1
/ GENERAL INFORMATION:
/ APPLICANT: Margarete Odenthal and Diana Jung
/ TITLE OF INVENTION: Gene Expression, Genome Alteration And Reporter Expression
/ FILE REFERENCE: 1472/68806
/ CURRENT APPLICATION NUMBER: US/10/332,733
/ PRIOR FILING DATE: 2003-06-10
/ NUMBER OF SEQ ID NOS: 41
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 22
/ LENGTH: 666
/ TYPE: DNA
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Description of artificial sequence: Red
/ OTHER INFORMATION: Fluorescent Protein
US-10-332-733-22

Alignment Scores:
Pred. No.: 1,256-140 Length: 666
Score: 1196.00 Matches: 221
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.52% Indels: 0
DB: 7 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-332-733-22 (1-666)
QY 5 LySAnVal1leLyegLuphMeArpPhelybVal1ArgMetGluG1YThRVal1Ang1Y 24
DB 1 AAGAATGTATCAAGAGGTTATGAGTTTAAAGTTTGATGAAGAAAGCGTCAAGTGAAGG 60
QY 25 H1SG1UPhEG1U1leG1UG1YGLUG1YARGProTYRG1UG1YH1SANG1Th1RYAL 44
```

```

Db      61 CACGAGTTTGAATAGAAAGCGAGGAGGAGGCGCATACAGAGGCCACATACCGTA 120
QY      45 LysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAsp1LeuSerProGln 64
Db      121 AGCTTAAAGTAAACCAAGGGGGGACCTTTGCGCATTTGGGATATTTTTCACACACA 180
QY      65 PheGlnTYrG1SerLysValTYrValLysHisProAlaAsp1LeuProAspTYrLysLys 84
Db      181 TTTCAAGTATGAAAGCAAGGTATATGTCAAGCACTTCGCACTACCAACTATATAAAG 240
QY      85 LeuSerPheProGlnGlyPheLysTrpGluArgValMetAsnPheGluAspGlyGlyVal 104
Db      241 CCGTCACTTCCCGAAGGATTTAAATGGAAAGGGTCATGAACTTTGAACGCGTGGCGTC 300
QY      105 ValThrValThrGlnAspSerSerLeuGlnAspGlyCysPhe1LeuLysValLysPhe 124
Db      301 GTTACTGTAAACCAAGATTCAGATTGACAGATGGCTGTTCATCTAACAGGTCAAGTTC 360
QY      125 T1EG1yValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTrpGlu 144
Db      361 ATTGGCGTAACTTTCTTCCTTCGATGACCTGTTATGCAAAAGAACATGGGCTGGGAA 420
QY      145 AlaSerThrGluArgLeuTYrProArgAspGlyValLeuLysGlyGluLeuHisLysVala 164
Db      421 GCCAGACTGAGCGGTTGTATCCTCGTATGCGGTGTGAAAGAGAGATTCATTAAGGCT 480
QY      165 LeuLysLeuLysAspGlyGlyHisTYrLeuValGluPheLysSer1LeuThrMetAlaLys 184
Db      481 CTGAAGCTGAAAGACGGTGTCTATCACTTAAGTGAATTCAAAAGTATTAATGGCAAG 540
QY      185 LysProValGlnLeuProGlyTYrTYrTYrValAspSerLysLeuAsp1LeuThrSerHis 204
Db      541 AAGCTGTGCACTACCAAGGTACTATGTTGACTCCAACTGGATTAACAACCCAC 600
QY      205 AsnGluAspTYrThr1LeuValGlnGlnTYrGluArgThrGluArgHisLysLeuPhe 224
Db      601 AACGAAGACTATACATCGTTGAGCAGTATGAAGAAGACGAGGACCATCTGTTCC 660
QY      225 Leu 225
Db      661 CTT 663

RESULT 48
US-10-314-936-1
; Sequence 1, Application US/10314936
; Publication No. US20040110225A1
; GENERAL INFORMATION:
; APPLICANT: Gibbs, Patrick D.L.
; APPLICANT: Schmale, Michael C.
; APPLICANT: Carter, Robert W.
; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES
; FILE REFERENCE: 638,004
; CURRENT APPLICATION NUMBER: US/10/314,936
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: mutant red fluorescent protein
; NAME/KEY: CDS
; LOCATION: (1)..(711)
; OTHER INFORMATION:
US-10-314-936-1

Alignment Scores:
Pred. No.: 1,37e-140 Length: 711
Score: 1196.00 Matches: 221
Percent Similarity: 99.11% Conservative: 2

```

```

Best Local Similarity: 98.22% Mismatches: 2
Query Match: 98.52% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-314-936-1 (1-711)

QY      1 MetArgSerSerLysAsnVal1LeuGluPheMetArgPheLysValArgMetGlnGly 20
Db      1 ATGAGTTGTCCAAATAATTATCAAGAGCTTCATGAGGTTTAAAGTTCAATGAGAGA 60
QY      21 ThrValAsnGlyHisGluPheGlu1LeuGluGlyGluGlyValArgProTYrGluGly 40
Db      61 ACGGTCAATGGGCACAGATTGAAATAGAAAGCGGAAGAGGAAGGACCAATCAAGAGCC 120
QY      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAsp1Le 60
Db      121 CACAATACCGTAAACCTTAAGGTAAACAAAGGGGGGACCTTTGCCATTTGCTTGGGATAT 180
QY      61 LeuSerProGlnPheGlnTYrG1SerLysValTYrValLysHisProAlaAsp1LePro 80
Db      181 TTGTCAACCAATATTCAGTATGGAAGCAAGATATGTCAAGCATCTCGGACATACCA 240
QY      81 AspTYrLysLysLeuSerPheProGlnGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db      241 GACTATTAAGGCTGTCTTCCTGAGAGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300
QY      101 AspGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPhe1LeuTYr 120
Db      301 GACGGTGGGCTGTTACTGTAAACCAAGATTCAGATTGCAAGATGGCTGTTCACTAC 360
QY      121 LysValLysPhe1LeuGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db      361 AAGTCAGATTCAATGGGCTGTAACCTTCTCTGATGACCTGTTATGCAAAAAGAGACA 420
QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTYrProArgAspGlyValLeuLysGlyGlu 160
Db      421 ATGGCTGGGAAGCCAGACCTGAGCGTTGTATCTCTGATAGGCGGTGTGAAGAGAG 480
QY      161 HisLysValaLeuLysLeuLysAspGlyGlyHisTYrLeuValGlnPheLysSer1Le 180
Db      481 ATTCATTAAGGCTCTAAGTTGAAGACGGTGTCTATTAAGTTGAATTCAAAACATATT 540
QY      181 TyrMetAlaLysLysProValGlnLeuProGlyTYrTYrTYrValAspSerLysLeuAsp 200
Db      541 TACATGGCAAGAAACCTGTGACCTACCAAGGTACTATGTTGACTCCAAACGTGAT 600
QY      201 1LeuThrSerHisAsnGluAspTYrThr1LeuValGlnGlnTYrGluArgThrGluArg 220
Db      601 ATTAACAAGCCCAACAAAGACTATCAATCGTTGAGCAGTATGAAGAAGACGAGGAGCC 660
QY      221 HisHisLysLeuPheLeu 225
Db      661 CACCACTGTCTCCT 675

RESULT 49
US-10-314-936-3
; Sequence 3, Application US/10314936
; Publication No. US20040110225A1
; GENERAL INFORMATION:
; APPLICANT: Gibbs, Patrick D.L.
; APPLICANT: Carter, Robert W.
; APPLICANT: Schmale, Michael C.
; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES
; FILE REFERENCE: 638,004
; CURRENT APPLICATION NUMBER: US/10/314,936
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:

```



OTHER INFORMATION: mutant red fluorescent protein  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(711)  
 OTHER INFORMATION:  
 US-10-314-936-3

## Alignment Scores:

Pred. No.:	1,37e-140	Length:	711
Score:	1196.00	Matches:	221
Percent Similarity:	99.11%	Conservative:	2
Best Local Similarity:	98.22%	Mismatches:	2
Query Match:	98.52%	Indels:	0
DB:	7	Gaps:	0

US-10-006-922a-12 (1-225) x US-10-314-936-3 (1-711)

```

QY 1 MetArgSerSerlybAenVal11lelybGluPhMeArGpHeLybValArgMeGluGly 20
   |||||
Db 1 ATGAGTGTTCACAAAGATGTTATCAAGAGTTCAAGAGTTTAAAGTTTCTGATGAAAGGA 60

QY 21 ThrValAsnGlyHleGluPhMeGlu1leGluGlyGluGlyArgProTyrgGluGly 40
   |||||
Db 61 ACGGTCAATGGGCGACGAGTTTGAATATGAAGCGAAGAGAGAGAGCGCATACGAAGGC 120

QY 41 HisAsnThrVallybLeuLybValThrLybGlyGlyProLeuProPhealatrPaaplle 60
   |||||
Db 121 CACAAATACCGTAAAGCTTAAGTTAACCAAGGGGGGACCTTGGCATTGCTGGGATATT 180

QY 61 LeuSerProGluPhMeGluTyrgLybSerLybValTyrgVallybHisProAlaapllePro 80
   |||||
Db 181 TTGTACACCAAAATTCAGTATGAGAACAGATATATCAAGCATCTGCGCAGATACCA 240

QY 81 AspGlyTybLybLeuSerPheProGluGlyPheLybTrrGluArgValMetAenPheGlu 100
   |||||
Db 241 GACTATATAAAGCTGTCAATTCCTGAAGATTTAAATGGAAAGGTCATGAACCTTGAA 300

QY 101 AspGlyGlyValValThrValThrGlnApsSerSerleuGlnAapGlyCybPhe1leTy 120
   |||||
Db 301 GACGTGCGCGTTCATGTTAACCAGAGATTCAGTTTGCGAGATGGCTGTTTCATCTAC 360

QY 121 LybVallybPhe1leGlyValAenPheProSerAapGlyProValMetGlnLybLybThr 140
   |||||
Db 361 AAGGTCAAGTTCATGCGTGAACCTTCTCTGATGACCTGTTATGACAAAGAACACA 420

QY 141 MetGlyTrrGluAlaSerThrGluArgLeuTyrgProArgAapGlyValleuLybGlyGlu 160
   |||||
Db 421 ATGGGCTGGGAAGCGACGACTGACGCTTGTATCTCTGATGACGCTGTTGAAGAGAG 480

QY 161 IleHisLybAlaLeuLybLeuLybAapGlyGlyHisTyrgLeuValGluPheLybSer1le 180
   |||||
Db 481 ATTCAATAAGGCTCGAAGTTGAAGACGGTGTCTATTACCTAATTGAATTCAAAACTATT 540

QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrgTyrgValAapSerLybLeuAap 200
   |||||
Db 541 TACATGGCAAGAACCTGTGCACTACAGGGTACTACTATGTGATCTCCAAACTGGAT 600

QY 201 IleHisSerHisAenGluAapTyrgThr1leValGluGlnTyrgLybArgThrGluGlyArg 220
   |||||
Db 601 ATAAACAAGCCCAACAAGACTATACATCGTTGAGCAGTATGAAGAACCAGAGAGAGC 660

QY 221 HisHisLeuPheLeu 225
   |||||
Db 661 CACCATCTGTTCTT 675
  
```

## RESULT 50

US-11-021-014-1  
 Sequence 1, Application US/11021014  
 Publication No. US20050100954A1  
 GENERAL INFORMATION:  
 APPLICANT: Gibbs, Patrick D.L.  
 APPLICANT: Carter, Robert W.  
 APPLICANT: Schmale, Michael C.

TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
 FILE REFERENCE: 639, 004  
 CURRENT APPLICATION NUMBER: US/11/021,014  
 CURRENT FILING DATE: 2004-12-23  
 PRIOR APPLICATION NUMBER: US/10/314,936  
 PRIOR FILING DATE: 2002-12-09  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 1  
 LENGTH: 711  
 TYPE: DNA  
 ORGANISM: Artificial  
 FEATURE:  
 OTHER INFORMATION: mutant red fluorescent protein  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(711)  
 OTHER INFORMATION:  
 US-11-021-014-1

## Alignment Scores:

Pred. No.:	1,37e-140	Length:	711
Score:	1196.00	Matches:	221
Percent Similarity:	99.11%	Conservative:	2
Best Local Similarity:	98.22%	Mismatches:	2
Query Match:	98.52%	Indels:	0
DB:	10	Gaps:	0

US-10-006-922a-12 (1-225) x US-11-021-014-1 (1-711)

```

QY 1 MetArgSerSerlybAenVal11lelybGluPhMeArGpHeLybValArgMeGluGly 20
   |||||
Db 1 ATGAGTGTTCACAAAGATGTTATCAAGAGTTTAAAGTTTAAAGTTTCTGATGAAAGGA 60

QY 21 ThrValAsnGlyHleGluPhMeGlu1leGluGlyGluGlyArgProTyrgGluGly 40
   |||||
Db 61 ACGGTCAATGGGCGACGAGTTTGAATATGAAGCGAAGAGAGAGAGCGCATACGAAGGC 120

QY 41 HisAsnThrVallybLeuLybValThrLybGlyGlyProLeuProPhealatrPaaplle 60
   |||||
Db 61 ACGGTCAATGGGCGACGAGTTTGAATATGAAGCGAAGAGAGAGAGCGCATACGAAGGC 120

QY 61 LeuSerProGluPhMeGluTyrgLybSerLybValTyrgVallybHisProAlaapllePro 80
   |||||
Db 181 TTGTACACCAAAATTCAGTATGAGAACAGATATATCAAGCATCTGCGCAGATACCA 240

QY 81 AspGlyTybLybLeuSerPheProGluGlyPheLybTrrGluArgValMetAenPheGlu 100
   |||||
Db 241 GACTATATAAAGCTGTCAATTCCTGAAGATTTAAATGGAAAGGTCATGAACCTTGAA 300

QY 101 AspGlyGlyValValThrValThrGlnApsSerSerleuGlnAapGlyCybPhe1leTy 120
   |||||
Db 301 GACGTGCGCGTTCATGTTAACCAGAGATTCAGTTTGCGAGATGGCTGTTTCATCTAC 360

QY 121 LybVallybPhe1leGlyValAenPheProSerAapGlyProValMetGlnLybLybThr 140
   |||||
Db 361 AAGGTCAAGTTCATGCGTGAACCTTCTCTGATGACCTGTTATGACAAAGAACACA 420

QY 141 MetGlyTrrGluAlaSerThrGluArgLeuTyrgProArgAapGlyValleuLybGlyGlu 160
   |||||
Db 421 ATGGGCTGGGAAGCGACGACTGACGCTTGTATCTCTGATGACGCTGTTGAAGAGAG 480

QY 161 IleHisLybAlaLeuLybLeuLybAapGlyGlyHisTyrgLeuValGluPheLybSer1le 180
   |||||
Db 481 ATTCAATAAGGCTCGAAGTTGAAGACGGTGTCTATTACCTAATTGAATTCAAAACTATT 540

QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrgTyrgValAapSerLybLeuAap 200
   |||||
Db 541 TACATGGCAAGAACCTGTGCACTACAGGGTACTACTATGTGATCTCCAAACTGGAT 600

QY 201 IleHisSerHisAenGluAapTyrgThr1leValGluGlnTyrgLybArgThrGluGlyArg 220
   |||||
Db 601 ATAAACAAGCCCAACAAGACTATACATCGTTGAGCAGTATGAAGAACCAGAGAGAGC 660
  
```

```

OY      221 H18H18LeuPheLeu 225
        |||||
DB      661 CACCATCTGTCTT 675

RESULT 51
US-11-021-014-3
; Sequence 3, Application US/11021014
; Publication No. US2005010954A1
GENERAL INFORMATION:
APPLICANT: Gibbs, Patrick D.L.
APPLICANT: Carter, Robert W.
APPLICANT: Schmale, Michael C.
TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES
FILE REFERENCE: 638.004
CURRENT APPLICATION NUMBER: US/11/021,014
CURRENT FILING DATE: 2004-12-23
PRIOR APPLICATION NUMBER: US/10/314,936
PRIOR FILING DATE: 2002-12-09
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patencin version 3.1
SEQ ID NO 3
LENGTH: 711
TYPE: DNA
ORGANISM: Artificial
FEATURES:
OTHER INFORMATION: mutant red fluorescent protein
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(711)
OTHER INFORMATION:
US-11-021-014-3

```

Alignment Scores:	
Pred. No.:	1,37e-140
Score:	119e-00
Percent Similarity:	99.11%
Best Local Similarity:	98.42%
Query Match:	98.52%
DB:	10
Length:	711
Matches:	222
Conservative:	2
Mismatches:	2
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) x US-11-021-014-3 (1-711)

QY	1	MetArgSerSerItyAsnValIleIyGcluphMetArgPheIyValArgMetGluGly	20
Db	1	ATGAGTTGTTCCAAAGATGTTATCAAGGAGTTCAAGAGTTAAAGTTCTGATGAAAGGA	60
QY	21	ThrValAsnGlyHfegIuphegiuilegiuGlyGluGlyValArgProTyGluGly	40
Db	61	ACGGTCATGGGACACAGTTTGAAATAGAGGCGAAGAGGAGGAGGCCATTACAAAGGC	120
QY	41	HisAsnThrValIyIeLeuIyValThrlYgIyGlyProLeuProPheAlaTPDAspIle	60
Db	121	CACAAATCCGTAAGCTTAAGTAAGTAACAAAGGGGGACCTTGTCATTGCTTGGGAAATT	180
QY	61	LeuSerProGluPheGluIntYrgIySerIyValIyValIyVhIleProAlaAspIlePro	80
Db	181	TTGTGCACCAAAATTCAGTATGGAGCAAGATATGTCAAGACTCTCGCGACATACCA	240
QY	81	AspTyrIleIyIeLeuSerPheProGluGlyPheIySTrPGluArgValMetAsnPheGlu	100
Db	241	GACTATATAAAAGCTGTCATTTCTCTGAAGATTTAAATGGGAAAGGCTCATAACTTTGAA	300
QY	101	AspGlyGlyValValaThrValThrgIuAspSerSerLeuGluAspGlyCysPheIleTyr	120
Db	301	GACGGTGGCCTCTTACTGTAAACCAAGAAATTCAGTTTGACGAGAGGCTGTTTCACTTAC	360
QY	121	IyValIyIyPheIleGlyValAsnPheProSerAspGlyProValMetGluIyIySerThr	140
Db	361	AAGGTCAAGTTCAATGGCGGAACCTTCTTCTGATGACCTGTTATGCAAAAGAACCA	420
QY	141	MetGlyTyrPGluAlaSerThrgIuArgLeuIyProArgAspGlyValIleuIyGlyGlu	160

Db	421	ATGGCGTGGAAAGCCACGACTGAGCGTTTGATCTCTGTATGGCGGTGTGAAGAAGAG	488
Qy	161	IIEHISLYEALALEULYSEULYSAPGLYGIYHSITYREUVALGIUPHELYSERILE	180
Db	481	ATTCAATTAAGGCTCTGAAAGTTGAAGAAGCGGTGCATTAACCTAAGTTGAATCAAACAATTT	540
Qy	181	TYMEEALILYLYBYPVOVALGINLEUPROGLIYYRTTYRVFVLAASPSELYBLEUASP	200
Db	541	TACATGGCAAGAGCGTCGTGCACTACCAAGSGRCTACTATGTGTGACTCCAACTGGAT	600
Qy	201	IIETHSERHISASNGIUNAPPTYRTHRIILEVALGIUGINTRYGIUARTRHGUGIYARG	220
Db	601	ATAACAGGCCAACAAAGACTATATCATTCGTTGACGATGAAAAGAACGAGGAGCACC	660
Qy	221	HIEHISLEUPHELEU	225
Db	661	CACCATCTGTCTCTT	675

RESULT 52  
US-10-006

US-10-006-922-36  
; Sequence 36, Application US/10006922  
; Publication No. US20020197676A1

```

1  GENERAL INFORMATION:
2  APPLICANT: Lukyanov, Sergey A
3  APPLICANT: Fradkov, Arcady F.
4  APPLICANT: Labas, Yulii A.
5  APPLICANT: Matz, Mikhail V.
6  APPLICANT: Tersikh, Alexey
7  TITLE OF INVENTION: No. US00020197676A1 Chromophores/Fluorophores and
8  TITLE OF INVENTION: Methods for Using the Same
9  FILE REFERENCE: CLON-035CIP
10 CURRENT APPLICATION NUMBER: US/10/006,922

```

? PRIOR APPLICATION NUMBER: 09/120,330  
 ? PRIOR FILING DATE: 1998-12-11  
 ? PRIOR APPLICATION NUMBER: 09/457,898  
 ? PRIOR FILING DATE: 1999-12-09  
 ? PRIOR APPLICATION NUMBER: 09/458,144  
 ? PRIOR FILING DATE: 1999-12-09  
 ? PRIOR APPLICATION NUMBER: 09/458,477

PRIOR APPLICATION NUMBER: 09/457,556

```

? PRIOR APPLICATION NUMBER: 09/444,338
? PRIOR FILING DATE: 1998-11-19
? NUMBER OF SEQ ID NOS: 46
? SOFTWARE: FastSeq for Windows Version 4.0.
? SEQ ID NO: 36
? LENGTH: 678
? TYPE: DNA
? ORGANISM: Dicosoma species
? US-10-006-922-36

```

**Alignment Scores:**

Pred. No.:	5 47e-140	Length:	678
Score:	1151.00	Matches:	220
Percent Similarity:	98.67%	Conservative:	2
Best Local Similarity:	97.78%	Mismatches:	3
Query Match:	98.11%	Indels:	0
DB:	5	Gaps:	0

US-10-006-922A-12 (1-225) x US-10-006-922-36 (1-678)

QY MeArgSerSerIysAsnValIleIysGluIubHekeArgPheIysValArgMetGluGly 20  
Db 1 ATGCGCTCTCTCCGAGAAAGCTCATCAACGAGTTCAATGCGCTTCAAGAGTGGCATGGAAGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTAAACGGCCACGAGTTCCAGATTCAGAGGCCAGAGGCCGCCCTTCAGAGAGGC 120  
QY HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTTPAspIle 60  
41

```

DB      121 CACAAACCGTGAAGTTGAAGGTGACCAAGGCGCCCTTCGCTCGGACATC 180
QY      61 LeuserProgl1nphedIntYrglyserLyValYrVallysh1aPro1aasp1lepro 80
DB      181 CTGTCCCCCAAGTTCAGTACGAGCTCCAGGCTGTACGTAGAGACCCCGCCGACATCCCC 240
QY      81 AapTYrLyalyseuSerPheProgluglyPheLyFTpGluArgYalMetanPheglu 100
DB      241 GACTACAAAGAGCTGCTTCCCGGAGGCTTCAGTGGGAGCGCGTGTGAATTCGAG 300
QY      101 AapGlyGlyValValThrValThrglnaapSerSerleuglnaapGlyCysPheileTy 120
DB      301 GACGGGGGGGCGAGCCCGTGAACCCAGACTCTCCCTGAGAGAGAGAGAGAGAGAGAG 360
QY      121 LyVallyPhe1leGlyVal1aenPheProSeraspGlyProValMetGlnlyslYthr 140
DB      361 MAGTGAAGTTCATCGGCGGTGAACTTCCTCCGAGGCGCGCGTGTATGCAAGAGAGAC 420
QY      141 MetGlyTYrGlnu1aSerThrglnaArgLeuTYrProArgaapGlyValleuLysglyGlu 160
DB      421 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGCGGAG 480
QY      161 l1eH1eLyAlaleuLyseuLyasaapGlyGlyH1eTYrleuValGlnPheLyser1le 180
DB      481 ATCCACAAAGCCCTGAAGCTGAAGAGACGGCGCACTGCTGTGAGATTCAAGTCCATC 540
QY      181 TYrMetAlaLyalySProValGlnLeuProGlyTYrTYrTYrVala1aPserLyseuasp 200
DB      541 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACACCAAGAGCTGAC 600
QY      201 l1eThSerH1aangluAapTYrThrl1eValGlnGlnIntYrgluArgThrgluArg 220
DB      601 ATCACTCCCAACAAGAGACTACACCATGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY      221 H1eH1eLeuPheLeu 225
DB      661 CACCACCTGTTCTG 675

RESULT 53
US-10-081-864-14
/ Sequence 14, Application US/10081864
/ Publication No. US2003022287A1
/ GENERAL INFORMATION:
/ APPLICANT: Lukyanov, Sergey
/ APPLICANT: Yanushevich, Yuriy
/ APPLICANT: Savitsky, Alexander
/ APPLICANT: Fradkov, Arcady
/ TITLE OF INVENTION: No. US2003022287A1 Aggregating Fluorescent Proteins and
/ FILE REFERENCE: Clon-067
/ CURRENT APPLICATION NUMBER: US/10/081,864
/ CURRENT FILING DATE: 2002-06-19
/ PRIOR APPLICATION NUMBER: 10/006,922
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 60/270,983
/ PRIOR FILING DATE: 2001-02-21
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: non-aggregating mutant
US-10-081-864-14

Alignment Scores:
Pred. No.: 5,47e-140 Length: 678
Score: 1191.00 Matches: 220
Percent Similarity: 98.67% Conservative: 2
Best Local Similarity: 97.78% Mismatches: 3
Query Match: 98.11% Indels: 0

```

```

DB:      5 Gaps: 0
US-10-006-922a-12 (1-225) x US-10-081-864-14 (1-678)
QY      1 MetArgSerSerLyshAnaVal1leLysgluPheMetArgPheLyValArgMetGluGly 20
DB      1 ATGGGCTCTCCGAGAGAGCTGATCACCGAGTTCATGCGCTTCAAGAGTGGCATGAGGGGC 60
QY      21 ThrValaanglyH1eGluPheglu1leGluGlyGluGlyGluGlyArgProTYrGlyGly 40
DB      61 ACCGTGAACGGGCACACAGTTCAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY      41 H1aenThVal1yLeuLyseuValThrlYsglyGlyProLeuProPhe1a1rPaap1le 60
DB      121 CACAAACCGTGAAGTTGAAGGTGACCAAGGCGCCCTTCGCTCGGACATC 180
QY      61 LeuserProgl1nphedIntYrglyserLyValYrVallysh1aPro1aasp1lepro 80
DB      181 CTGTCCCCCAAGTTCAGTACGAGCTCCAGGCTGTACGTAGAGACCCCGCCGACATCCCC 240
QY      81 AapTYrLyalyseuSerPheProgluglyPheLyFTpGluArgYalMetanPheglu 100
DB      241 GACTACAAAGAGCTGCTTCCCGGAGGCTTCAGTGGGAGCGCGTGTGAATTCGAG 300
QY      101 AapGlyGlyValValThrValThrglnaapSerSerleuglnaapGlyCysPheileTy 120
DB      301 GACGGGGGGGCGAGCCCGTGAACCCAGACTCTCCCTGAGAGAGAGAGAGAGAGAGAG 360
QY      121 LyVallyPhe1leGlyVal1aenPheProSeraspGlyProValMetGlnlyslYthr 140
DB      361 MAGTGAAGTTCATCGGCGGTGAACTTCCTCCGAGGCGCGCGTGTATGCAAGAGAGAC 420
QY      141 MetGlyTYrGlnu1aSerThrglnaArgLeuTYrProArgaapGlyValleuLysglyGlu 160
DB      421 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGCGGAG 480
QY      161 l1eH1eLyAlaleuLyseuLyasaapGlyGlyH1eTYrleuValGlnPheLyser1le 180
DB      481 ATCCACAAAGCCCTGAAGCTGAAGAGACGGCGCACTGCTGTGAGATTCAAGTCCATC 540
QY      181 TYrMetAlaLyalySProValGlnLeuProGlyTYrTYrTYrVala1aPserLyseuasp 200
DB      541 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACACCAAGAGCTGAC 600
QY      201 l1eThSerH1aangluAapTYrThrl1eValGlnGlnIntYrgluArgThrgluArg 220
DB      601 ATCACTCCCAACAAGAGACTACACCATGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY      221 H1eH1eLeuPheLeu 225
DB      661 CACCACCTGTTCTG 675

RESULT 54
US-10-315-920-5
/ Sequence 5, Application US/10315920
/ Publication No. US20030175809A1
/ GENERAL INFORMATION:
/ APPLICANT: Tersikh, Alexey
/ APPLICANT: Fradkov, Arcady Fedorovich
/ TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
/ FILE REFERENCE: Clon-077CIP
/ CURRENT APPLICATION NUMBER: US/10/315,920
/ CURRENT FILING DATE: 2002-12-09
/ PRIOR APPLICATION NUMBER: 60/211,607
/ PRIOR FILING DATE: 2000-06-14
/ PRIOR APPLICATION NUMBER: PCT/US01/19097
/ PRIOR FILING DATE: 2001-06-13
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 678
/ TYPE: DNA

```

ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: variant of sequence from *Discosoma* sp.  
 US-10-315-920-5

## Alignment Scores:

Pred. No.:	5,47e-140	Length:	678
Score:	1191.00	Matches:	220
Percent Similarity:	98.67%	Conservative:	2
Best Local Similarity:	97.78%	Mismatches:	3
Query Match:	98.11%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922a-12 (1-225) x US-10-315-920-5 (1-678)

```

QY      1 MetAAGSerSerLyAaenValIleLySGluPhMeArpPhelyValAArgMetGluGly 20
DB      1 ATGGCCTCTCCGAGAACGTCATCACCGAGTTCAATCGCTTCAAGGTGGCATGAGGGGC 60
QY      21 ThrValAaNGlYHlAgLuPhGluIleGluGluGluGluGluGluGluGluGluGluGluGlu 40
DB      61 ACCGTGAACGCCACAGTTCAGATCGAGGGCGAGGGCGAGGGCGGCGCTTACAGAGGC 120
QY      41 HisAenThrValLyLeuLyValThrlYsgLyGlyProLeuProPhaLaATpAapIle 60
DB      121 CAACAACCGTGAAGCTGAAGGTGACCAAGGGGGCGGCGGCGCTTGGCTGGGACATC 180
QY      61 LeuSerProGluPhGluInTrGlySerLyValYrValLyHlAProAaapIlePro 80
DB      181 CTGTCCCCCAGATTCAGATCGAGGTTCACAGGTGATGACCAAGGTGATGACCAAGGTGATGAC 240
QY      81 AApTYrLybLyLeuSerPheProGluGluGluGluGluGluGluGluGluGluGluGluGlu 100
DB      241 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGGCGCTGATGAATTCGAG 300
QY      101 AApGlyGlyValValThrlValThrgInAaPserSerLeuGlnAaPglyCyapPheIleYr 120
DB      301 GACGGGGGGGTGGCGACCGTGACCGAGACTCTCCCTCGACGAGCGGCTTCAATCTAC 360
QY      121 LybValLybPheIleGlyValAaenPheProSerAaPglyProValMetGluLybYrThr 140
DB      361 AAGGTGAATTATCGGCTGAACTTCCCTCCGACGGCGCGGTGATGAGAAAGAAC 420
QY      141 MetGlyTrpGluAaSerThrgInuArgLeuYrProAaArgAaPglyValLeuLybGlyGlu 160
DB      421 ATGGGCTGGAGGCTTCCACCGAGGCTGTACCCCGGAGCGGCTGTGAAGGGCGAG 480
QY      161 IleHlAlybAlaLeuLybLeuLybAaPglyGlyHlAlyrLeuValGluPhelybSerIle 180
DB      481 ATCCCAAGCGCCCTGAAGCTGAAGAGCGGGGCACTACCTGTGGAGTTCAAGTCAATC 540
QY      181 TyrMetAlaLybLybProValGlnLeuPProGlyYrYrYrYrValAaPserLybLeuAaP 200
DB      541 TACATGGCCCAAGAGCCCGTGAAGCTGCGCGGCTACTACTACGTGACACCAAGCTGGAGC 600
QY      201 IleThrSerHisAaNGluAaPTrThrlIleValGluGlnYrGluAaArgHrGluGluAaG 220
DB      601 ATCACTCCCAACAAGAGCTACACATCATGTGAGAGCAATGAGAGCGACCGAGGGCGGC 660
QY      221 HisHlAlybPheLeu 225
DB      661 CACCACTGTCTCTG 675

```

RESULT 55  
 US-10-006-922-38  
 ; Sequence 38, Application US/10006922  
 ; Publication No. US20020197676A1

GENERAL INFORMATION:  
 ; APPLICANT: Lukyanov, Sergey A  
 ; APPLICANT: Fradkov, Arcady F.  
 ; APPLICANT: Labas, Yulii A.  
 ; APPLICANT: Matz, Mikhail V.  
 ; APPLICANT: Terelikh, Aleksey

TITLE OF INVENTION: No. US20020197676A1 Chromophores/fluorophores and  
 ; TITLE OF INVENTION: Methods for using the same  
 ; FILE REFERENCE: CLON-035CTP  
 ; CURRENT APPLICATION NUMBER: US/10/006,922  
 ; PRIOR FILING DATE: 2001-12-04  
 ; PRIOR APPLICATION NUMBER: 09/120,330  
 ; PRIOR FILING DATE: 1998-12-11  
 ; PRIOR APPLICATION NUMBER: 09/457,898  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: 09/458,144  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: 09/458,477  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: 09/457,556  
 ; PRIOR FILING DATE: 1999-12-09  
 ; PRIOR APPLICATION NUMBER: 09/444,338  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 38  
 ; LENGTH: 675  
 ; TYPE: DNA  
 ; ORGANISM: *Discosoma* species  
 ; US-10-006-922-38

## Alignment Scores:

Pred. No.:	2,32e-139	Length:	675
Score:	1186.00	Matches:	219
Percent Similarity:	98.22%	Conservative:	2
Best Local Similarity:	97.33%	Mismatches:	4
Query Match:	97.69%	Indels:	0
DB:	5	Gaps:	0

US-10-006-922a-12 (1-225) x US-10-006-922-38 (1-675)

```

QY      1 MetAAGSerSerLyAaenValIleLySGluPhMeArpPhelyValAArgMetGluGly 20
DB      1 ATGGCCTCTCCGAGAACGTCATCACCGAGTTCAATCGCTTCAAGGTGGCATGAGGGGC 60
QY      21 ThrValAaNGlYHlAgLuPhGluIleGluGluGluGluGluGluGluGluGluGluGluGlu 40
DB      61 ACCGTGAACGCCACAGTTCAGATCGAGGGCGAGGGCGAGGGCGGCGCTTACAGAGGC 120
QY      41 HisAenThrValLyLeuLyValThrlYsgLyGlyProLeuProPhaLaATpAapIle 60
DB      121 CAACAACCGTGAAGCTGAAGGTGACCAAGGGGGCGGCGGCGCTTGGCTGGGACATC 180
QY      61 LeuSerProGluPhGluInTrGlySerLyValYrValLyHlAProAaapIlePro 80
DB      181 CTGTCCCCCAGATTCAGATCGAGGTTCACAGGTGATGACCAAGGTGATGACCAAGGTGATGAC 240
QY      81 AApTYrLybLyLeuSerPheProGluGluGluGluGluGluGluGluGluGluGluGluGlu 100
DB      241 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGGCGGTGATGAATTCGAG 300
QY      101 AApGlyGlyValValThrlValThrgInAaPserSerLeuGlnAaPglyCyapPheIleYr 120
DB      301 GACGGGGGGGTGGCGACCGTGACCGAGACTCTCCCTCGACGAGCGGCTTCAATCTAC 360
QY      121 LybValLybPheIleGlyValAaenPheProSerAaPglyProValMetGluLybYrThr 140
DB      361 AAGGTGAATTATCGGCTGAACTTCCCTCCGACGGCGCGGTGATGAGAAAGAAC 420
QY      141 MetGlyTrpGluAaSerThrgInuArgLeuYrProAaArgAaPglyValLeuLybGlyGlu 160
DB      421 ATGGGCTGGAGGCTTCCACCGAGGCTGTACCCCGGAGCGGCTGTGAAGGGCGAG 480
QY      161 IleHlAlybAlaLeuLybLeuLybAaPglyGlyHlAlyrLeuValGluPhelybSerIle 180
DB      481 ATCCCAAGCGCCCTGAAGCTGAAGAGCGGGGCACTACCTGTGGAGTTCAAGTCAATC 540
QY      181 TyrMetAlaLybLybProValGlnLeuPProGlyYrYrYrYrValAaPserLybLeuAaP 200

```

Db	541	TACATGCGCCAAAGAACCCCTGCAAGCTGCCCCGGCATCTACTACGTAAGGACCCCAAGCTGGAC	600
QY	201	IIlethSerHisangluAapTYrThIleValgluIntYrgluAqThrgluYarg	220
Db	601	ATACACCTCCCAACAGAGCATACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCGCG	660
QY	221	HisHisleuPhelen	225
Db	661	CACCACTGTTCTG	675
RESULT 56			
US-10-081-864-13			
Sequence 13, Application US/10081864			
Publication No. US20030022287A1			
GENERAL INFORMATION:			
APPLICANT: Lukyanov, Sergey			
APPLICANT: Yanyushevich, Yuriy			
APPLICANT: Savitsky, Alexander			
APPLICANT: Fedkov, Arcady			
TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and			
FILE REFERENCE: CLON-067			
CURRENT APPLICATION NUMBER: US/10/081, 864			
CURRENT FILING DATE: 2002-06-19			
PRIOR APPLICATION NUMBER: 10/006, 922			
PRIOR FILING DATE: 2001-12-04			
PRIOR APPLICATION NUMBER: 60/270, 983			
PRIOR FILING DATE: 2001-02-21			
NUMBER OF SEQ ID NOS: 30			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 13			
LENGTH: 675			
TYPE: DNA			
ORGANISM: Artificial Sequence			
FEATURE:			
OTHER INFORMATION: non-aggregating mutant			
US-10-081-864-13			
Alignment Scores:			
Pred. No.: 2,32e-139 Length: 675			
Score: 1186.00 Matches: 219			
Percent Similarity: 98.22% Conservative: 2			
Best Local Similarity: 97.33% Mismatches: 4			
Query Match: 97.69% Indels: 0			
DB: 5 Gaps: 0			
US-10-006-922A-12 (1-225) x US-10-081-864-13 (1-675)			
QY	1	MetArgSerSerIyBAAnValIIleYgluPhenMetArgPhelyValArgMetgluY	20
Db	1	ATGGCCTCTCTCGAGAACGTCAATCAACCAAGTTCATGCCCTTCAGGTGCAATGAGGCG	60
QY	21	ThrValAangIyHIsgluPhelgluIIlegluYgluYgluYargProTYrgluY	40
Db	61	ACCGTGAACGGCCACGAGTTTCAGAGATCGAGGGCGAGGGCGGGCCCTTACGAGAGGCG	120
QY	41	HisAnthrValYleYleuYValThryYglYProleuProPhelAATrPhaPile	60
Db	121	CACAACACCGGTGAAGCTGAAGGTGACCAAGGGCGGGCCCTTCCCTTGGCTGGGACATC	180
QY	61	LeuSerProgluInpHeglInTYrgIySerIyValTYrValIyHIsProIlaAspIlePro	80
Db	181	CTGTCCCCCAATCTCAAGTACGAGCTTCAGAGTGTACGTGAGACACCCCGCAKATCCCC	240
QY	81	AspTYrYleYleuSerPheProgluYPhelySTrPGIuArgValMetAnpHeglIu	100
Db	241	GACTCAAGAAAGCTGTCTTCCCGGAGGCTTCAAGTGGAGACGGGTGATGAATTCGAG	300
QY	101	AspGlyglYValValThrValThrgInAspSerSerIeuGInAspGlyCyapHelleTYr	120
Db	301	GACGGCGGGGTGGCGACCGGTGACCGAGGCTCTCCCTGCAAGACAGCGGCTTCAATCAG	360

QY	121	LYVALYbPhe11eGlyValaAnPheProSerAspGlyProValMetGlnIlySerThr	140
Db	361	AAAGTGAAGTTCAATCGCGCGGAATCTCCCTCCGAGCGCCCGTGAAGCAAGAAAGACC	420
QY	141	MetG1TTrG1u1aSerThrG1uArGLeuTYrProArGAspG1YValIeuysG1yG1u	160
Db	421	ATGGGCTGGAGAGGCTTCACCGAGCCCTGTACCCCGGAGCGCGGTGTGAAGGGCGAG	480
QY	161	11eH1sYalAleuIlySeuIysAspG1yG1yH1sTYrLeuValG1uPheIySerIle	180
Db	481	ACCACAAAGCCCTGAAGCTGAAGAGACGGCGCCACTACTGTGTGAAGTTCAAGTCCATC	540
QY	181	TYrMetAlaIyIysProValG1uLeuProG1TYrTYrTYrValaAspSerIySLeuAsp	200
Db	541	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAAGCGCAAGCTGCAC	600
QY	201	11eThSerH1sAnG1uAepTYrThrIleValaG1uG1uTYrG1uArGThrG1uG1yArG	220
Db	601	ATCACCTCCCAACAAGAGACTACACCATGTGTGAGAGTACGAGCGCACCGAGGGCCGC	660
QY	221	H1eH1sLeuPheIeu 225	
Db	661	CACCACTGTCTCTG 675	

  

QY	121	LYVALYbPhe11eGlyValaAnPheProSerAspGlyProValMetGlnIlySerThr	140
Db	361	AAAGTGAAGTTCAATCGCGCGGAATCTCCCTCCGAGCGCCCGTGAAGCAAGAAAGACC	420
QY	141	MetG1TTrG1u1aSerThrG1uArGLeuTYrProArGAspG1YValIeuysG1yG1u	160
Db	421	ATGGGCTGGAGAGGCTTCACCGAGCCCTGTACCCCGGAGCGCGGTGTGAAGGGCGAG	480
QY	161	11eH1sYalAleuIlySeuIysAspG1yG1yH1sTYrLeuValG1uPheIySerIle	180
Db	481	ACCACAAAGCCCTGAAGCTGAAGAGACGGCGCCACTACTGTGTGAAGTTCAAGTCCATC	540
QY	181	TYrMetAlaIyIysProValG1uLeuProG1TYrTYrTYrValaAspSerIySLeuAsp	200
Db	541	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAAGCGCAAGCTGCAC	600
QY	201	11eThSerH1sAnG1uAepTYrThrIleValaG1uG1uTYrG1uArGThrG1uG1yArG	220
Db	601	ATCACCTCCCAACAAGAGACTACACCATGTGTGAGAGTACGAGCGCACCGAGGGCCGC	660
QY	221	H1eH1sLeuPheIeu 225	
Db	661	CACCACTGTCTCTG 675	

  

QY	121	LYVALYbPhe11eGlyValaAnPheProSerAspGlyProValMetGlnIlySerThr	140
Db	361	AAAGTGAAGTTCAATCGCGCGGAATCTCCCTCCGAGCGCCCGTGAAGCAAGAAAGACC	420
QY	141	MetG1TTrG1u1aSerThrG1uArGLeuTYrProArGAspG1YValIeuysG1yG1u	160
Db	421	ATGGGCTGGAGAGGCTTCACCGAGCCCTGTACCCCGGAGCGCGGTGTGAAGGGCGAG	480
QY	161	11eH1sYalAleuIlySeuIysAspG1yG1yH1sTYrLeuValG1uPheIySerIle	180
Db	481	ACCACAAAGCCCTGAAGCTGAAGAGACGGCGCCACTACTGTGTGAAGTTCAAGTCCATC	540
QY	181	TYrMetAlaIyIysProValG1uLeuProG1TYrTYrTYrValaAspSerIySLeuAsp	200
Db	541	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAAGCGCAAGCTGCAC	600
QY	201	11eThSerH1sAnG1uAepTYrThrIleValaG1uG1uTYrG1uArGThrG1uG1yArG	220
Db	601	ATCACCTCCCAACAAGAGACTACACCATGTGTGAGAGTACGAGCGCACCGAGGGCCGC	660
QY	221	H1eH1sLeuPheIeu 225	
Db	661	CACCACTGTCTCTG 675	

  

QY	121	LYVALYbPhe11eGlyValaAnPheProSerAspGlyProValMetGlnIlySerThr	140
Db	361	AAAGTGAAGTTCAATCGCGCGGAATCTCCCTCCGAGCGCCCGTGAAGCAAGAAAGACC	420
QY	141	MetG1TTrG1u1aSerThrG1uArGLeuTYrProArGAspG1YValIeuysG1yG1u	160
Db	421	ATGGGCTGGAGAGGCTTCACCGAGCCCTGTACCCCGGAGCGCGGTGTGAAGGGCGAG	480
QY	161	11eH1sYalAleuIlySeuIysAspG1yG1yH1sTYrLeuValG1uPheIySerIle	180
Db	481	ACCACAAAGCCCTGAAGCTGAAGAGACGGCGCCACTACTGTGTGAAGTTCAAGTCCATC	540
QY	181	TYrMetAlaIyIysProValG1uLeuProG1TYrTYrTYrValaAspSerIySLeuAsp	200
Db	541	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAAGCGCAAGCTGCAC	600
QY	201	11eThSerH1sAnG1uAepTYrThrIleValaG1uG1uTYrG1uArGThrG1uG1yArG	220
Db	601	ATCACCTCCCAACAAGAGACTACACCATGTGTGAGAGTACGAGCGCACCGAGGGCCGC	660
QY	221	H1eH1sLeuPheIeu 225	
Db	661	CACCACTGTCTCTG 675	

  

QY	121	LYVALYbPhe11eGlyValaAnPheProSerAspGlyProValMetGlnIlySerThr	140
Db	361	AAAGTGAAGTTCAATCGCGCGGAATCTCCCTCCGAGCGCCCGTGAAGCAAGAAAGACC	420
QY	141	MetG1TTrG1u1aSerThrG1uArGLeuTYrProArGAspG1YValIeuysG1yG1u	160
Db	421	ATGGGCTGGAGAGGCTTCACCGAGCCCTGTACCCCGGAGCGCGGTGTGAAGGGCGAG	480
QY	161	11eH1sYalAleuIlySeuIysAspG1yG1yH1sTYrLeuValG1uPheIySerIle	180
Db	481	ACCACAAAGCCCTGAAGCTGAAGAGACGGCGCCACTACTGTGTGAAGTTCAAGTCCATC	540
QY	181	TYrMetAlaIyIysProValG1uLeuProG1TYrTYrTYrValaAspSerIySLeuAsp	200
Db	541	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAAGCGCAAGCTGCAC	600
QY	201	11eThSerH1sAnG1uAepTYrThrIleValaG1uG1uTYrG1uArGThrG1uG1yArG	220
Db	601	ATCACCTCCCAACAAGAGACTACACCATGTGTGAGAGTACGAGCGCACCGAGGGCCGC	660
QY	221	H1eH1sLeuPheIeu 225	
Db	661	CACCACTGTCTCTG 675	

  

QY	121	LYVALYbPhe11eGlyValaAnPheProSerAspGlyProValMetGlnIlySerThr	140
Db	361	AAAGTGAAGTTCAATCGCGCGGAATCTCCCTCCGAGCGCCCGTGAAGCAAGAAAGACC	420
QY	141	MetG	





```
FILE REFERENCE: US 1353/03 (VA)
CURRENT APPLICATION NUMBER: US/10/742,828
CURRENT FILING DATE: 2003-12-23
PRIOR APPLICATION NUMBER: US 10/306,737
PRIOR FILING DATE: 2002-11-29
NUMBER OF SEQ ID NOS: 5
SEQ ID NO 5
LENGTH: 7495
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: Unsure
LOCATION:
OTHER INFORMATION: Synthesized
US-10-742-828-5

Alignment Scores:
Pred. No.: 6,646-138 Length: 7495
Score: 1186.00 Matches: 219
Percent Similarity: 98.22% Conservative: 2
Best Local Similarity: 97.33% Mismatches: 4
Query Match: 97.69% Indels: 0
DB: Gaps: 7

US-10-006-922a-12 (1-225) x US-10-742-828-5 (1-7495)

QY 1 MetAAGSerSerLyAsnValIlelysgIuphMeCArgPhelyValArgMetGluGly 20
DB 2123 ATGGGCTCTCCGAGAAAGCTCATACCGAGTTCAAGCTTCAAGGTGGCATGAGAGGC 2182
QY 21 ThrValAenGlyHhIegIuphEgIuilegIuGlyGluGlyArgProTyrgIuGly 40
DB 2103 ACCGGAAACGGCAAGATTGCAAGATCGAGGCGAGGCGGAGGCGCCCTACAGAGGC 2242
QY 41 HhAsnThrValIleuLyValThrysgIyProLeuProPheAlaTrpAspIle 60
DB 2243 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTCCCTCGGAGCATATC 2302
QY 61 LeuSerProGlnPhEgInTrGlySerLyValIlyValIyHhAsProAlaAspIlePro 80
DB 2303 CTGTCCCCCAAGTTCAGAGCTCCAGGTGATCGTGAAGCAACCCGCGACATCCCC 2362
QY 81 AspTyrlLyLyLeuSerPheProGluGlyPheLySTrPGluArgValMetAsnPhEgIn 100
DB 2363 GACTACAAAGACTGTCTCTCCCGAGGCTTCAAGTGGAGCGGCTGATGAATTCGAG 2422
QY 101 AspGlyGlyValIValThryValThrgInAspSerSerLeuGlnAspGlyCysePheIleTy 120
DB 2423 GACGGCGGCGTGGCGACCGGTGACCCGAGACTCTCTCGAGAGCGGCTTCAATCTAC 2482
QY 121 LysValIlyAspPheIleGlyValIAsnPhEProSerAspGlyProValMetGlnIylysthr 140
DB 2483 AAGGTGAAGTTATCGGCTGAACTTCCCTCCGAGGCGCGGTGATGAGAAAGAAC 2542
QY 141 MetGlyTrPGluAlaSerThrGluArgLeuTyProArgAspGlyValIleuLysgIyGlu 160
DB 2543 ATGGCTGGGAAGCTCTCCACCGAGCCCTGTACCCCGGAGCGGCTGTGAAGGGCGAG 2602
QY 161 IleHhIyValIleuLyLeuLyAspGlyGlyHhIyTrLeuValGluPhelysSerIle 180
DB 2603 ACCCAAGAGCCCTGAAGCTGAAGGACGGCGGCGCACTACCTGTGAGATTCAAGTCTATC 2662
QY 181 TyrMetAlaIylyLyProValGlnIyProGlyTyrTyrTrValIAspSerLyseuAsp 200
DB 2663 TACATGGCCCAAGAGCCCTGAGCTGCGCGGCTACTACTAGTGAACCGCAAGCTGAGAC 2722
QY 201 IleThrSerHhAsnGlyuAspTyrlThryIleValGluGlnTyArgThrGlyGluGlyArg 220
DB 2723 ATCACTCTCCCAAGAGGACTACACATCGTGAAGAGTACGAGCGCACCGAGGGCCGC 2782
QY 221 HhIhIleuPhelyu 225
DB 2783 CACCACTGTCTCTG 2797
```

```
RESULT 62
US-10-510-363-5
Sequence 5, Application US/10510363
Publication No. US20050158724A1
GENERAL INFORMATION:
APPLICANT: Yoder, Richard A.
TITLE OF INVENTION: METHODS OF IDENTIFYING COMPOUNDS THAT MODULATE A DNA REPAIR
TITLE OF INVENTION: PATHWAY AND/OR RETRO-VIRAL INFECTIVITY, THE COMPOUNDS, AND USES
FILE REFERENCE: SCHN-0033
CURRENT APPLICATION NUMBER: US/10/510,363
PRIOR FILING DATE: 2004-10-05
PRIOR APPLICATION NUMBER: PCT/US03/10302
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: US 60/370,376
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.3
SEQ ID NO 5
LENGTH: 9731
TYPE: DNA
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthesized retroviral vectors
US-10-510-363-5

Alignment Scores:
Pred. No.: 9,556-138 Length: 9731
Score: 1186.00 Matches: 219
Percent Similarity: 98.22% Conservative: 2
Best Local Similarity: 97.33% Mismatches: 4
Query Match: 97.69% Indels: 0
DB: Gaps: 9

US-10-006-922a-12 (1-225) x US-10-510-363-5 (1-9731)

QY 1 MetAAGSerSerLyAsnValIlelysgIuphMeCArgPhelyValArgMetGluGly 20
DB 5496 ATGGCTCTCCGAGAAAGCTCATACCGAGTTCAAGCTTCAAGGTGGCATGAGAGGC 5555
QY 21 ThrValAenGlyHhIegIuphEgIuilegIuGlyGluGlyArgProTyrgIuGly 40
DB 5556 ACCGGAAACGGCAAGATTGCAAGATCGAGGCGAGGCGGAGGCGCCCTACAGAGGC 5615
QY 41 HhAsnThrValIleuLyValThrysgIyProLeuProPheAlaTrpAspIle 60
DB 5616 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTCCCTCGGAGCATATC 5675
QY 61 LeuSerProGlnPhEgInTrGlySerLyValIlyValIyHhAsProAlaAspIlePro 80
DB 5676 CTGTCCCCCAAGTTCAGAGCTCCAGGTGATCGTGAAGCAACCCGCGACATCCCC 5735
QY 81 AspTyrlLyLyLeuSerPheProGluGlyPheLySTrPGluArgValMetAsnPhEgIn 100
DB 5736 GACTACAAAGACTGTCTCTCCCGAGGCTTCAAGTGGAGCGGCTGATGAATTCGAG 5795
QY 101 AspGlyGlyValIValThryValThrgInAspSerSerLeuGlnAspGlyCysePheIleTy 120
DB 5796 GACGGCGGCGTGGCGACCGGTGACCCGAGACTCTCTCGAGAGCGGCTTCAATCTAC 5855
QY 121 LysValIlyAspPheIleGlyValIAsnPhEProSerAspGlyProValMetGlnIylysthr 140
DB 5856 AAGGTGAAGTTATCGGCTGAACTTCCCTCCGAGGCGCGGTGATGAGAAAGAAC 5915
QY 141 MetGlyTrPGluAlaSerThrGluArgLeuTyProArgAspGlyValIleuLysgIyGlu 160
DB 5916 ATGGCTGGGAAGCTCTCCACCGAGCCCTGTACCCCGGAGCGGCTGTGAAGGGCGAG 5975
QY 161 IleHhIyValIleuLyLeuLyAspGlyGlyHhIyTrLeuValGluPhelysSerIle 180
DB 5976 ACCCAAGAGCCCTGAAGCTGAAGGACGGCGGCGCACTACCTGTGAGATTCAAGTCTATC 6035
```





QY 21 ThrValaenGIyHhIeGIuPhelGluIleGIuGIyGIuGIyAArgProTyrGIuGIy 40  
Db 153 AGGGTCMAATGGGCGACGAGTTTGAATAAGAAAGCGAAAGAGGAGGCGCATACGAAGGC 212  
QY 41 HisaenThrValIyLeuLeuValThrIySGIyGIyProLeuProPheAATrPApIle 60  
Db 213 CACMAATACCGTAAAGCTTAAGGTAAACCAAGGGGACCTTTGCCATTCTGGGATATT 272  
QY 61 LeuSerProGInPheGInTyrGIySerIyValIyTrValIyHisProAlaApIlePro 80  
Db 273 TTGTCAACCAATTTCAGATGAGAAAGCAAGATATATGTCAAGCAACCTGCCACATACCA 332  
QY 81 AApTyrIyIyLeuLeuSerPheProGIuGIyPhelYsTrPGIuArgValMetAsnPheGIu 100  
Db 333 GACTATAAAGAACTGTGATTCCTGAGGAGATTTAAATGGGAAAGGTCATVGAACCTTTGAA 392  
QY 101 AApGIyGIyValIyValThrValIyThrGInAsPSeSerLeuGInAspGIyCyAsPheIleTyr 120  
Db 393 GACGGTGGCGTCTTAATCTTAACCCAGATTCAGTTTGCAGATGGCTGTTTCATCTAC 452  
QY 121 IySValIyIyPheIleGIyValAsnPheProSerAspGIyProValMetGInIyIySThr 140  
Db 453 AAGTCMAAGTTCAATGGCTTGAACCTTCCCTCCGATGACCTGTATGCAAAAGAAAGACA 512  
QY 141 MetGIyTrPGIuAlaSerThrGIuArgLeuTyrProArgAspGIyValIleuIySGIyGIu 160  
Db 513 ATGGGCTGGGAAGCCAGCACTGAGCGCTTGTATCCCTCGATGGCGCTGTGAAAGGAGAG 572  
QY 161 IleHisIyAlaLeuIyLeuIyAspGIyGIyHisIyTrLeuValGIuPhelYsSerIle 180  
Db 573 ATTCATTAAGCTCTGAAGCTGAAAGCGGTGTCATTACTTAAGTTCAATTCAAAAGTATT 632  
QY 181 TyrMetAlaIyLeuIyAspProValGIuLeuProGIyTyrTyrTyrValAspSerIyLeuAsp 200  
Db 633 TACATGGCAAGAGAGCTGTGACAGCTACCAAGGTACTACTATGTGACCTCAAACTGAGAT 692  
QY 201 IleThrSerHisAsnGIuAspTyrThrIleValGIuGIuInTyrGIuArgThrGIuArg 220  
Db 693 ATTAACAAAGCAACAAACGAACTATACATCGTTGAGCAGTATGAAAGAAAGCAAGGAGCGC 752  
QY 221 HisHisLeuPheLeu 225  
Db 753 CACCACTCTTCTCTT 767  
RESULT 65  
US-10-121-258-5  
; Sequence 5, Application US/10121258  
; Publication No. US20030059835A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/10/121,258  
; CURRENT FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; US-10-121-258-5  
Alignment Scores:

Pred. No.: 4,476-136 Length: 678  
Score: 1160.00 Matches: 216  
Percent Similarity: 97.33% Conservative: 3  
Best Local Similarity: 96.00% Mismatches: 6  
Query Match: 95.55% Indels: 0  
DB: 5 Gaps: 0  
US-10-006-922a-12 (1-225) x US-10-121-258-5 (1-678)  
QY 1 MetAsPSeSerIyAsnValIleIySGIuPhelMetArgPheIyValArgMetGIuGIy 20  
Db 1 ATGGCTCTCCCGAGAGAGCTCATCAAGAGATTCAAGCCCTTCMAAGTGGCGATGAGGGCG 60  
QY 21 ThrValaenGIyHhIeGIuPhelGluIleGIuGIyGIuGIyAArgProTyrGIuGIy 40  
Db 61 TCCTGTAACGGCCACGAGATTGCAAGTCAAGAGCGAGGCGAGGCGCCCTTACGAGGC 120  
QY 41 HisaenThrValIyLeuLeuValThrIySGIyGIyProLeuProPheAATrPApIle 60  
Db 121 ACCAGACCGCCAGACTGAAGGTGACCAAGGCGGCGCCCTTGCTGGCGATC 180  
QY 61 LeuSerProGInPheGInTyrGIySerIyValIyTrValIyHisProAlaApIlePro 80  
Db 181 CTGTCCCCCAAGTTCAGATACGGCTCAAGGTGTACGTGAAGCAACCCGCCACATCCCC 240  
QY 81 AApTyrIyIyLeuLeuSerPheProGIuGIyPhelYsTrPGIuArgValMetAsnPheGIu 100  
Db 241 GACTCAAGAAAGCTCTCTCCCGAGGCTTCAAGTGGAGCGGCGGTATGAACTTCGAG 300  
QY 101 AApGIyGIyValIyValThrValIyThrGInAsPSeSerLeuGInAspGIyCyAsPheIleTyr 120  
Db 301 GACGGCGGGCGTGTGATACCGTGAACCAAGACTCTCCCTCGAGAGAGCGGTCTTCATCTAC 360  
QY 121 IySValIyIyPheIleGIyValAsnPheProSerAspGIyProValMetGInIyIySThr 140  
Db 361 AAGGTGAAGTTCAATGGGCTGTAACCTTCCCTCCGACGCGCCCGTAAATGCAAGAAAGACT 420  
QY 141 MetGIyTrPGIuAlaSerThrGIuArgLeuTyrProArgAspGIyValIleuIySGIyGIu 160  
Db 421 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGGAGCGGCGTGTGAAAGGCGAG 480  
QY 161 IleHisIyAlaLeuIyLeuIyAspGIyGIyHisIyTrLeuValGIuPhelYsSerIle 180  
Db 481 ATCCCAAGAGCCCTGAAGCTGAAGGACGGCGGCACCTACTGTGTGAAGTTCAAGTCCATC 540  
QY 181 TyrMetAlaIyLeuIyAspProValGIuLeuProGIyTyrTyrTyrValAspSerIyLeuAsp 200  
Db 541 TACATGGCCAAAGAACCCGTGCAAGCTGCCCGGCTACTACTAGTGAATCCAAAGCTGAGAC 600  
QY 201 IleThrSerHisAsnGIuAspTyrThrIleValGIuGIuInTyrGIuArgThrGIuArg 220  
Db 601 ATCACTCCCAACAAAGAGACTACACATCGTGAAGCAGTACAGCGCGCCGAGGGCGCC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACTCTTCTCTG 675  
RESULT 66  
US-10-931-304-5  
; Sequence 5, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258

```
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FASTSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
/ OTHER INFORMATION: "T1"
US-10-931-304-5

Alignment Scores:
Pred. No.: 4,47e-136 Length: 678
Score: 1160.00 Matches: 216
Percent Similarity: 97.33% Conservative: 3
Best Local Similarity: 96.00% Mismatches: 6
Query Match: 95.55% Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-5 (1-678)

QY 1 MetArgSerSerIysAsnValIleIysGluPhmeArGpHeIysValArgMetGluGly 20
DB 1 ATGGCTCTCTCCAGAGACGTCAACAGAGACTTATGGCTTCAAGAGGCGCATGAGGCG 60
QY 21 ThrValAsnGlyHISgluphegluIleGluGlyGluGlyGluGlyArgProIYrGluGly 40
DB 61 TCGGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGCGGCGGCGGCGGCGGCGGCGG 120
QY 41 HisAsnThrValIysLeuIysValIThrIysGlyGlyProIeuProPheIaITpAspIle 60
DB 121 ACCCAGCGCCCAAGCTGAAGGTGACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 61 LeuSerProGlnPheGlnIYrGlySerIysValIYrValIYsHISProIaAspIlePro 80
DB 181 CTGTCCCCCAAGTTCAGTACGGCTCCAGGTGACGTGAAGCAACCCCGCGCATCTCC 240
QY 81 AspTYrIlyIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100
DB 241 GACTACAAAGAGCTGTCTCTCCCGAGGCTTCAAGTGGAGCGCGGTGATGAACCTTCAG 300
QY 101 AspGlyGlyValValIThrValIThrGlnAspSerSerIeuGlnAspGlyCySpheIleTYr 120
DB 301 GACGGCGGCGGTGTGACCGTGAACCGAGACTCTCTCGCAAGAGGCGGCTCTCATCTAC 360
QY 121 IysValIysPheIleGlyValIAsnPheProSerAspGlyProValMetGlnIYsIYsThr 140
DB 361 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACCGGCGGCGGCGGCGGCGGCGG 420
QY 141 MetGlyTrpGluIAsnThrGluArgLeuTYrProIArgAspGlyValIleuIysGlyGlu 160
DB 421 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGGCGGTGAAGGCGGAG 480
QY 161 IleHisIysValIleuIysLeuIysAspGlyGlyHISTrIeuValIglupheIysSerIle 180
DB 481 ATCCAAAGGCGCTGAAGCTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 540
QY 181 TyMetAlaIysIysProValGlnLeuProGlyTYrTYrTYrValIAspSerIysLeuAsp 200
DB 541 TACATGGCCAAAGAGCCCGTGAAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY 201 IleThrSerHisAsnGluAspTYrThrIleValIgluGlnIYrGluArgThrGluGlyArg 220
DB 601 ATCACTCTCCCAACGAGACTACACCATCTGTGAGCAGTACGAGCGGCGGAGGCGGCGG 660
QY 221 HisHisIeuPheIeu 225
DB 661 CACCACTGTTCCTG 675
```

```
RESULT 67
US-10-844-064A-3
/ Sequence 3, Application US/10844064A
/ Publication No. US20050149994A1
/ GENERAL INFORMATION:
/ APPLICANT: Glisc, Benjamin
/ TITLE OF INVENTION: RAPIDLY MATURING FLUORESCENT PROTEINS AND METHODS FOR USING THE S
/ FILE REFERENCE: 092234-9006
/ CURRENT APPLICATION NUMBER: US/10/844,064A
/ PRIOR FILING DATE: 2004-05-11
/ PRIOR APPLICATION NUMBER: PCT/US02/40539
/ PRIOR FILING DATE: 2002-12-18
/ PRIOR APPLICATION NUMBER: US 60/341,723
/ PRIOR FILING DATE: 2001-12-19
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 3
/ LENGTH: 704
/ TYPE: DNA
/ ORGANISM: Drosophila
US-10-844-064A-3

Alignment Scores:
Pred. No.: 4,72e-136 Length: 704
Score: 1160.00 Matches: 216
Percent Similarity: 97.33% Conservative: 3
Best Local Similarity: 96.00% Mismatches: 6
Query Match: 95.55% Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-844-064A-3 (1-704)

QY 1 MetArgSerSerIysAsnValIleIysGluPhmeArGpHeIysValArgMetGluGly 20
DB 20 ATGGCTCTCTCCAGAGACGTCAACAGAGACTTATGGCTTCAAGAGGCGCATGAGGCG 79
QY 21 ThrValAsnGlyHISgluphegluIleGluGlyGluGlyGluGlyArgProIYrGluGly 40
DB 80 TCGGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGG 139
QY 41 HisAsnThrValIysLeuIysValIThrIysGlyGlyProIeuProPheIaITpAspIle 60
DB 140 ACCCAGCGCCCAAGCTGAAGGTGACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 199
QY 61 LeuSerProGlnPheGlnIYrGlySerIysValIYrValIYsHISProIaAspIlePro 80
DB 200 CTGTCCCCCAAGTTCAGTACGGCTCCAGGTGACGTGAAGCAACCCCGCGCATCTCC 259
QY 81 AspTYrIlyIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100
DB 260 GACTACAAAGAGCTGTCTCTCCCGAGGCTTCAAGTGGAGCGGCGGTGATGAACCTTCAG 319
QY 101 AspGlyGlyValValIThrValIThrGlnAspSerSerIeuGlnAspGlyCySpheIleTYr 120
DB 320 GACGGCGGCGGTGTGACCGTGAACCGAGACTCTCTCGCAAGAGGCGGCTCTCATCTAC 379
QY 121 IysValIysPheIleGlyValIAsnPheProSerAspGlyProValMetGlnIYsIYsThr 140
DB 380 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACCGGCGGCGGCGGCGGCGGCGG 439
QY 141 MetGlyTrpGluIAsnThrGluArgLeuTYrProIArgAspGlyValIleuIysGlyGlu 160
DB 440 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGGCGGTGAAGGCGGAG 499
QY 161 IleHisIysValIleuIysLeuIysAspGlyGlyHISTrIeuValIglupheIysSerIle 180
DB 500 ATCCAAAGGCGCTGAAGCTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 559
QY 181 TyMetAlaIysIysProValGlnLeuProGlyTYrTYrTYrValIAspSerIysLeuAsp 200
DB 560 TACATGGCCAAAGAGCCCGTGAAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 619
```





Percent Similarity: 95.96%  
Best Local Similarity: 92.83%  
Query Match: 10  
Conservative: 7  
Matches: 9  
Indels: 0  
Gaps: 0

US-10-006-922a-12 (1-225) x US-11-052-001-5 (1-1809)

```
QY 3 SerSerLyAsnValIleLygluPheMetArgPheLyValArgMetGluGlyThrVal 22
DB TCCTCGAGAGTCATCAAGAGTTCAAGCGCTTCAAGGTGAGAGGAGGCTCCG 483
QY 23 AsnGlyHISgluPheGluIleGluGlyGluGlyArgProGlyValArgMet 42
DB AACGCCCAAGATTGAGATCGAGGCGGAGGCGGCGGCGCTTCAAGAGGCA 543
QY 43 ThrValLyLeuLyValThrLyGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62
DB ACCGCCAAGCTGAAAGTGCACAAAGGCGGCGCTTCAAGCGGAGCATCTGTC 603
QY 63 ProGlnPheGlnTrpGlySerLyValThrValLyAsnProAlaAspIleProAspTrp 82
DB CCCCGTTCCAGTACGCTCCCAAGCGTACGTGAAGCACCCGCGACATCCCGA 663
QY 83 LySlyLeuSerPheProGluGlyPheLyStrpGluArgValMetAsnPheGluAspGly 102
DB AAGAGCTGCTCTCCCGAGGCGCTTCAAGTGGAGCGCGGTATGATGACCTGAG 723
QY 103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTrpVal 122
DB GCGCGTGCAGCCGTACCCAGGAGCTCTCCCTGCGAGAGCGCGCATGTTCAAG 783
QY 123 LySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLyLeuSerThrMetGly 142
DB AAGTTCGCGGACCAACTTCCCCCGAGGCGCGCTTCAAGAGAGAGACCATG 843
QY 143 TrpGluAlaSerThrGluArgLeuTrpProArgAspGlyValLeuLySgluIleHis 162
DB TCGAGAGCTCCACCGAGGCGCTGTAACCCCGCGAGCGGCTGTAAGGCGAG 903
QY 163 LySAlaLeuLyLeuLyAspGlyGlyHISTrpLeuValGlnPheLySerIleTrpMet 182
DB CAGCGCTGAAAGCTCAAGGACGCGCGCTTCAAGTGAAGCATCTTCAAG 963
QY 183 AlaLySlyProValGlnLeuProGlyTrpTrpValAspSerLyLeuAspIleThr 202
DB GCGAAGAGCCGTGACGCTGCGGCTACTAGTGAACCAAGCTGAGCATCAC 1023
QY 203 SerHisAsnGluAspTrpThrIleValGluGlnTrpGluArgThrGluGlyArgHis 222
DB TCCCAACAAGAGATCAACCATGTAAGAGAGCGCTCCGAGGCGCGCACAC 1083
QY 223 LeuPheLeu 225
DB 1084 CTGTTCTG 1092
```

RESUL 72  
US-10-006-922-43  
Sequence 43, Application US/10006922  
Publication No. US20020197676A1  
GENERAL INFORMATION:  
APPLICANT: Lukyanov, Sergey A  
APPLICANT: Pradkov, Arcady F.  
APPLICANT: Labas, Yulii A.  
APPLICANT: Matz, Mikhail V.  
APPLICANT: Teresikh, Aleksey  
TITLE OF INVENTION: No. US20020197676A1 Chromophores/Fluorophores and  
TITLE OF INVENTION: Methods for using the same  
FILE REFERENCE: C10N-035CTP  
CURRENT APPLICATION NUMBER: US/10/006,922  
CURRENT FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: 09/120,330  
PRIOR FILING DATE: 1998-12-11  
PRIOR APPLICATION NUMBER: 09/457,898

```
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/458,144
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/458,477
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/457,556
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/444,338
/ PRIOR FILING DATE: 1999-11-19
/ NUMBER OF SEQ ID NOS: 46
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO: 43
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: hybrid construct
US-10-006-922-43
```

Alignment Scores:  
Pred. No.: 678-131  
Score: 1119.00  
Length: 678  
Matches: 206  
Percent Similarity: 96.89%  
Conservative: 12  
Best Local Similarity: 91.56%  
Mismatches: 7  
Query Match: 92.17%  
Indels: 0  
Gaps: 0

US-10-006-922a-12 (1-225) x US-10-006-922-43 (1-678)

```
QY 1 MetArgSerSerLyAsnValIleLygluPheMetArgPheLyValArgMetGluGly 20
DB 1 ATGAGCTCAGACCAAGAAAGTATCAAGAGTTCATGCGGTTCAAGGTCCGATGAGG 60
QY 21 ThrValAsnGlyHISgluPheGluIleGluGlyGluGlyArgProGlyTrpGluGly 40
DB 61 ACCGGAAGCGGACAGATTGAGTCAAGGCGGAGGCGGCGGCGCTTCAAGAGG 120
QY 41 HisAsnThrValLyLeuLyValThrLySgluGlyProLeuProPheAlaTrpAspIle 60
DB 121 CACTGACAGCTGAACTCATGTGACCAAGGCGGCGCTTCCCTTCCCTTGCATC 180
QY 61 LeuSerProGlnPheGlnTrpGlySerLyValThrValLyAsnProAlaAspIlePro 80
DB 181 CTCAGCGCCCGATTCAATGAGGAGCAAGGTGTAAGTGAAGCACCCGCGCATCC 240
QY 81 AspTrpLySlyLeuSerPheProGluGlyPheLyStrpGluArgValMetAsnPheGlu 100
DB 241 GACTACAAAGAGCTCAAGCTTCCCGAGGCTTCAAGTGGAGCGGCTGATGAAC 300
QY 101 AspGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTrp 120
DB 301 GACGCGCGCTGTGATCCGTAGCCAGACAGACCTCAAGAGCGCTTCAATCAC 360
QY 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLyLeuThr 140
DB 361 GAGGTGAAGTTCATGCGGTGAATCTCCAGCGAGCGCGCTGATGACGCGGAGC 420
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTrpProArgAspGlyValLeuLySglu 160
DB 421 CGGCGCTGAGGCGGAGGAGGAGGCTTCAACCCCGAGCGGCTGCTCAAGGCGAG 480
QY 161 IleHisLySAlaLeuLyLeuLyAspGlyGlyHISTrpLeuValGlnPheLySerIle 180
DB 481 ATCCAGATGCGCTTCCGCTGAGGCGGCGGCGCATCTCTGTGAGTTCAGAGCATC 540
QY 181 TyrMetAlaLySlyProValGlnLeuProGlyTrpTrpValAspSerLyLeuAsp 200
DB 541 TACATGCGCAAGAGCCGTGAGCTCCCGGCTACTAGTGAACAGCAAGCTCGAC 600
QY 201 IleThrSerHisAsnGluAspTrpThrIleValGluGlnTrpGluArgThrGluGlyArg 220
DB 601 ATCACACAGCCCAAGAGATCAACCATGTGAGAGTGAAGAGCGGAGCGAGGCGG 660
```

QY 221 HSHLEUPHELU 225  
DB 661 CACCACTCTCCTC 675

RESULT 73  
US-10-081-864-11  
/ Sequence 11, Application US/10081864  
/ Publication No. US20030022287A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Lukyanov, Sergey  
/ APPLICANT: Kabanov, Konstantin  
/ APPLICANT: Yambuevich, Yuriy  
/ APPLICANT: Savitsky, Alexandr  
/ APPLICANT: Pradkov, Arcady  
/ TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and  
/ TITLE OF INVENTION: Methods for Using the Same  
/ FILE REFERENCE: CLON-067  
/ CURRENT APPLICATION NUMBER: US/10/081,864  
/ CURRENT FILING DATE: 2002-06-19  
/ PRIOR APPLICATION NUMBER: 10/006,922  
/ PRIOR FILING DATE: 2001-12-04  
/ PRIOR APPLICATION NUMBER: 60/270,983  
/ PRIOR FILING DATE: 2001-02-21  
/ NUMBER OF SEQ ID NOS: 30  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 11  
/ LENGTH: 678  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ OTHER INFORMATION: hybrid coding sequence  
US-10-081-864-11

Alignment Scores:  
Pred. No.: 6.7e-131 Length: 678  
Score: 1119.00 Matches: 206  
Percent Similarity: 96.89% Conservative: 12  
Best Local Similarity: 91.56% Mismatches: 7  
Query Match: 92.17% Indels: 0  
Gaps: 0

US-10-006-922a-12 (1-225) x US-10-081-864-11 (1-678)

QY 1 MetArgSerSerIysaenValIlelysgluPhemeArpPheylsValArgMetgluY 20  
DB 1 ATAGAGCTGCAGCAAGACGTGATCAAGAGTTCAAGGTTCAAGGCGGAGTGAAGGC 60  
QY 21 ThrValaenglyHtsgluphegluilegluglygluglygluYArgProtyrYrYgluY 40  
DB 61 ACCGTGACGGCCGACGAGTTCGAGATCAAGGCGGCGGCGGCGGCGCTTACGAGGGC 120  
QY 41 HtsaenThrValIysleuYsValThrlyglYglYProleuProPhealATPAspIle 60  
DB 121 CACTGCAGCGGTGAGCTCATGTGTACCAAGGCGGCGGCGGCGGCGGCGGCTTGCATC 180  
QY 61 leuSerProglInpHehInTYrGlySerIysValTYrValIysHisProAlaAspIlePro 80  
DB 181 CTGAGCCCCCAAGTTCCAGTACGACGACGAGGTGACCTGAAAGCACCCCGCGACATCCCC 240  
QY 81 AspTYrIysIysleuSerPheProgluYlYPhelysTYrPgluArgValMetAsnPhelglu 100  
DB 241 GACACCAAGAAAGCTCAGCTTCCCGAGGGCTTCAAGTGGGAGCGGGTGAATGAATTCGAG 300  
QY 101 AspglyglYValIThrValThrGlnaPseSerleuGlnaAspGlyCysePheIleTYr 120  
DB 301 GACGCGCGCGGTGAGCGTGAAGCAGGACGAGCTCAAGGAGCGGCTGCTTCACTCAC 360  
QY 121 LysValIysPheIleGlyValAsnPharPseSerAspGlyProValMetGlnIysleYThr 140  
DB 361 GAGGTGAAGTTCACTCGGCTGAATCTCCCGACGACGCGCGCGGTGATGACGGCGGAGCC 420  
QY 141 MetGlyTYrPgluAlaSerThrGluArgleuTYrProArgAspGlyValIleuYglYglu 160

DB 421 CGGGGTGGAGAGCCAGCAGCGGCGCTTACCCCCGGAGCGGGCTGCTCAAGGCGAC 480  
QY 161 ILeHslysaIaleuYsleuYsAspGlyYHtsYrIleuValGlnPhelysSerIle 180  
DB 481 ATCCACATGGCCCTCCGCTCGAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540  
QY 181 TyrMetAlaIysleuYsProValGlnleuPProglYTYrTYrValAspSerIysleuAsp 200  
DB 541 TACATGGCCAAAGAGCCCGTGAAGTCTCCCGGCTACTACTAGTGAACAGAGCTCGAC 600  
QY 201 ILeThrSerHisAsnGluAspTYrThrIleValIgluInTYrGluArgThrGluArg 220  
DB 601 ATCAGCAGCCACAGAGAGACTACACATCGTGAAGTGAAGAGGAGGAGGAGGCGCG 660  
QY 221 HSHLEUPHELU 225  
DB 661 CACCACTCTCCTC 675

RESULT 74  
US-10-931-304-82  
/ Sequence 82, Application US/10931304  
/ Publication No. US20050196768A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Tsien, Roger  
/ APPLICANT: Campbell, Robert  
/ APPLICANT: Baird, Geoffrey  
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
/ FILE REFERENCE: 39754-0831CP2CP3  
/ CURRENT APPLICATION NUMBER: US/10/931,304  
/ CURRENT FILING DATE: 2004-08-30  
/ PRIOR APPLICATION NUMBER: 10/209,208  
/ PRIOR FILING DATE: 2002-07-29  
/ PRIOR APPLICATION NUMBER: 10/121,258  
/ PRIOR FILING DATE: 2002-04-10  
/ PRIOR APPLICATION NUMBER: 09/866,538  
/ PRIOR FILING DATE: 2001-05-24  
/ PRIOR APPLICATION NUMBER: 09/794,308  
/ PRIOR FILING DATE: 2001-02-26  
/ NUMBER OF SEQ ID NOS: 110  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 82  
/ LENGTH: 705  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ OTHER INFORMATION: Polynuc. of Polypeptide variant "dimer2,2MM(dimer3) (dtomato)"  
US-10-931-304-82

Alignment Scores:  
Pred. No.: 7.57e-128 Length: 705  
Score: 1095.00 Matches: 201  
Percent Similarity: 94.57% Conservative: 8  
Best Local Similarity: 90.95% Mismatches: 12  
Query Match: 90.20% Indels: 0  
Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-82 (1-705)

QY 5 LysaenValIlelysgluPhemeArpPheylsValArgMetgluYThrValaengly 24  
DB 16 GAGAGGTGATCAAGAGTTCATGCTTCAAGGTGAGTGAAGGCGGCTCATAAGAGCGC 75  
QY 25 HtsgluphegluilegluglygluYArgProtyrYrYgluYHtsaenThrVal 44  
DB 76 CACGAGTTCAGATCGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 135  
QY 45 LysleuYsValThrIysglYglYProleuProPhealATPAspIleleuSerProglIn 64  
DB 136 AAGGTGAAGTGAACAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 195  
QY 65 PheGlnTYrGlySerIysleuYsValTYrValIysHisProAlaAspIleProAspTYrYsleY 84

DB 196 TTGATGTAAGCGCTCAAGGCGTACGTGAAGACACCCCGCCGACATCCCGATTACAAGAG 255  
QY 85 LeuSerPheProGluGlyPheIleTyrTrpGluValMetAsnPheGluAspGlyGlyVal 104  
DB 256 CTGTCCTTCCCGAGGCTTCAAGTGGAGCGCGTATCAACTTCGAGGACGCGGCTCG 315  
QY 105 ValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyrIleValIlePhe 124  
DB 316 GTGACCTGATACCCAGACCTCTCTCCGTGACGAGCGACCGCTGATCTACAGGTGAAGAG 375  
QY 125 IlegIValAsnPheProSerAspGlyProValMetGluIleValThrMetGlyTyrGlu 144  
DB 376 CGCGGACCAACTTCCCGCCGAGCGCCCGTATGACAGAGACCATGGGCTGGAG 435  
QY 145 AlaSerThrGluArgLeuTyrProArgAspGlyValIleuIleGlyGluIleHisVala 164  
DB 436 GCCTCCACCGAGCGCTGTATCCCGCCGACGCGGTCTAAGGGCGAGATCCACCGAGCC 495  
QY 165 LeuIleValAspGlyValIleIleTyrLeuValGluPheIleSerIleTyrMetAlaIys 184  
DB 496 CTGAAGCTAAGAGACGCGCGCCACTCTGCTGAGATTCAAGACCATCTGAGCCAG 555  
QY 185 IysProValGluLeuProGlyTyrTyrTyrValAspSerIleLeuAspIleThrSerHis 204  
DB 556 AAGCCCGTCAACTGCGCGGCTACTACTAGTGAACACCAAGCTGACATCCCTCCAC 615  
QY 205 AengIAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisIlePhe 224  
DB 616 AACGAGACTACACCATCTGTGAACGATGAGCGCTCGAGGGCGCCACCACTGTTCC 675  
QY 225 Leu 225  
DB 676 CTG 678  
RESULT 75  
US-10-931-107  
; Sequence 107, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; PRIOR FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 107  
; LENGTH: 1431  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "tdTomato"  
US-10-931-304-107  
Alignment Scores:  
Pseq. No.: 2,03e-127 Length: 1431  
Score: 1095.00 Matches: 201  
Percent Similarity: 94.57% Conservative: 8  
Best Local Similarity: 90.95% Mismatches: 12  
Query Match: 90.20% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-107 (1-1431)  
QY 5 LysAsnValIleIleValGluPheMetArgPheIleValArgMetGluGlyThrValAengIy 24  
DB 16 GAGGAGTCATCAAGAGATTCATGCGCTTCAAGGTGCGCATGAGAGGCTCCCATGAACGGC 75  
QY 25 HisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsnThrVal 44  
DB 76 CACGAGTTGAGATGAGAGGCGAGGCGAGGCGCGCCCTTACGAGAGGACCCAGACCGCC 135  
QY 45 LysLeuIysValThrLysGlyGlyProLeuProPheAlaTyrPaspIleLeuSerProGln 64  
DB 136 AAGCTGAAGTGAACCAAGGCGCGCCCTGCGCTTCCGCTGGAGATCTGTCCTCCCGCCAG 195  
QY 65 PheGlnTyrGlySerIleValTyrValIleHisProAlaAspGlyIleProAspTyrIleIys 84  
DB 196 TTGATGTAAGCGCTCAAGGCGGTATGTAAGCACCCCGCGAGATCCCGCATTAACAAGAG 255  
QY 85 LeuSerPheProGluGlyPheIleTyrTrpGluValMetAsnPheGluAspGlyGlyVal 104  
DB 256 CTGTCCTTCCCGAGGCTTCAAGTGGAGCGCGGTGATGAATCTGAGGACGCGGCTCG 315  
QY 105 ValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyrIleValIlePhe 124  
DB 316 GTGACCTGATACCCAGACCTCTCTCCGTGACGAGCGCACGCTGATCTACAGGTGAAGAG 375  
QY 125 IlegIValAsnPheProSerAspGlyProValMetGluIleValThrMetGlyTyrGlu 144  
DB 376 CGCGGACCAACTTCCCGCCGAGCGCCCGTATGACAGAGAACACCATGGGCTGGAG 435  
QY 145 AlaSerThrGluArgLeuTyrProArgAspGlyValIleuIleGlyGluIleHisVala 164  
DB 436 GCCTCCACCGAGCGCTGTATCCCGCCGACGCGGTCTGAAGGGGAGATCCACCGAGCC 495  
QY 165 LeuIleValAspGlyValIleIleTyrLeuValGluPheIleSerIleTyrMetAlaIys 184  
DB 496 CTGAAGCTGAAGAGACGCGCGCCACTACTGCTGAGATTCAAGACCATCTGAGCCAG 555  
QY 185 IysProValGluLeuProGlyTyrTyrTyrValAspSerIleLeuAspIleThrSerHis 204  
DB 556 AAGCCCGTCAACTGCGCGGCTACTACTAGTGAACACCAAGCTGACATCCCTCCAC 615  
QY 205 AengIAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisIlePhe 224  
DB 616 AACGAGACTACACCATCTGTGAACGATGAGCGCTCGAGGGCGCCACCACTGTTCC 675  
QY 225 Leu 225  
DB 676 CTG 678  
RESULT 76  
US-10-006-922-17  
; Sequence 17, Application US/10006922  
; Publication No. US20020197676A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey A  
; APPLICANT: Fradkov, Arcady F.  
; APPLICANT: Labas, Yulii A.  
; APPLICANT: Matz, Mikhail V.  
; APPLICANT: Terekikh, Aleksey  
; TITLE OF INVENTION: No. US20020197676A1 Chromophores/Fluorophores and  
; TITLE OF INVENTION: Methods for using the same  
; FILE REFERENCE: CLON-035C1P  
; CURRENT APPLICATION NUMBER: US/10/006,922  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 09/120,330  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/457,898  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,144  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,477  
; PRIOR FILING DATE: 1999-12-09



```
/ PRIOR APPLICATION NUMBER: 09/457,556
/ PRIOR FILING DATE: 1999-12-09
/ PRIOR APPLICATION NUMBER: 09/444,338
/ PRIOR FILING DATE: 1999-11-19
/ NUMBER OF SEQ ID NOS: 46
/ SOFTWARE: PairedSeq for Windows Version 4.0
/ SEQ ID NO 17
/ LENGTH: 876
/ TYPE: DNA
/ ORGANISM: Diacosoma species
US-10-006-922-17

Alignment Scores:
Pred. No.: 1,62e-126 Length: 876
Score: 1085.50 Matches: 199
Percent Similarity: 95.58% Conservative: 17
Best Local Similarity: 88.05% Mismatches: 9
Query Match: 89.42% Indels: 1
DB: Gaps: 1

US-10-006-922a-12 (1-225) x US-10-006-922-17 (1-876)
QY 1 MetArSerSerlysaenValIlelyGluPheMetArPheLyValAArgMetGluGly 20
DB 45 ATGAGTTGTTCCAAAGATGATCAAGAGATTCATGAGTTCAAGTTCCATATGAAAGCA 104
QY 21 ThValAAnGlyHtHegluPheGluIleGluGlyGluGlyAArgProTyrgLugly 40
DB 105 AGGTCATATGGCAAGAGTTGAAATTAAGGCAAGAGTGAAGGAGCCTTCAAGAGGT 164
QY 41 HisAAnThVallyLeuLyValThThLyGlyGlyProLeuProPheAlaTPAPhile 60
DB 165 CACGTTCGCTAAAGCTTATGATGATCAAGAGTGAAGTTCCTTCATTTGATATAT 224
QY 61 LeuSerProGluPheGluTyrgLySerLyValTyrrVallyshHsPheAlaAPhilePro 80
DB 225 TTGTCAACACAAATTCATGATGAAAGCAAGGATATATGCAACCCCTGCCGACATACCA 284
QY 81 AspTyrtLyshLyLeuSerPhePheProGluGlyPheLystrPgiuAArgValMetAAnPheGlu 100
DB 285 GACATATTAAGTCAATTCCTTCCTGAGGATTTAAATGGAAAGGTCATGAATTTGAA 344
QY 101 AspGlyGlyValValThThValThThGlnApsSerLeuGlnApsGlyCyshPheIleTyrr 120
DB 345 GAGGTCGCGGTGATCTGATATCCCAAGATTCAGTTGAAGACGGCTGTTTCATCTAC 404
QY 121 LyshVallyshPheIleGlyValAAnPheProSerApsGlyProValMetGlnLyshLyThr 140
DB 405 GAGGTCAAGTTCATTTGGGGTGAACCTTCCTTGATGACCTGTTATGACAGAGAGACA 464
QY 141 MetGlyTrpGluAlaSerThThGluArgLeuTyrrProAArgApsGlyValIleuLyGlyGlu 160
DB 465 CGGGGCTGGGAAGCCAGCTCTGACGCTTTGATCTCTGATATGGGGTCTGAAAGAGAC 524
QY 161 IleHsHsVallyLeuLyValApsGlyGlyLyshLyLeuValGluPheLySerIle 180
DB 525 ATCATATGGCTCTGAGAGCTGGAAGAGCGGCATTAACCTGTTGAATTCAAAAGATAT 584
QY 181 TyrrMetAlaLyshLyPro--ValGlnLeuProGlyTyrrTyrrValApsSerLyLeu 199
DB 585 TACATGTGTAAGAACCTTCAGTCAGATGGCAAGCTTACTATATATGTTGACTTCAAACTG 644
QY 200 AspIleThrSerHsAenGluApsTyrrThrIleValGluGluTyrgLuglyAArgThrgLugly 219
DB 645 GATATGACAGCCCAACAGAAATTAACAAGTCTGTGACGATGATAAAAAACCAGGGA 704
QY 220 ArgHsHsLeuPheLeu 225
DB 705 CGCCACCATTCGTTCAAT 722

RESULT 77
US-10-161-403-39
/ Sequence 39, Application US/10161403
```

```
/ Publication No. US20030119104A1
/ GENERAL INFORMATION:
/ APPLICANT: Perkins, Edward
/ APPLICANT: Perez, Carl
/ APPLICANT: Lindabaum, Michael
/ APPLICANT: Greene, Amy
/ APPLICANT: Leung, Josephine
/ APPLICANT: Fleming, Elena
/ APPLICANT: Stewart, Sandra
/ APPLICANT: Shellard, Joan
/ TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
/ FILE REFERENCE: 24601-420
/ CURRENT APPLICATION NUMBER: US/10/161,403
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: 60/294,758
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: 60/366,891
/ PRIOR FILING DATE: 2002-03-21
/ NUMBER OF SEQ ID NOS: 129
/ SOFTWARE: PairedSeq for Windows Version 4.0
/ SEQ ID NO 39
/ LENGTH: 876
/ TYPE: DNA
/ ORGANISM: Diacosoma species
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (45)...(737)
/ OTHER INFORMATION: Nucleotide sequence encoding red fluorescent
/ OTHER INFORMATION: protein (PP593)
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: GenBank AF272711
/ DATABASE ENTRY DATE: 2000-09-26
US-10-161-403-39

Alignment Scores:
Pred. No.: 1,62e-126 Length: 876
Score: 1085.50 Matches: 199
Percent Similarity: 95.58% Conservative: 17
Best Local Similarity: 88.05% Mismatches: 9
Query Match: 89.42% Indels: 1
DB: Gaps: 1

US-10-006-922a-12 (1-225) x US-10-161-403-39 (1-876)
QY 1 MetArSerSerlysaenValIlelyGluPheMetArPheLyValAArgMetGluGly 20
DB 45 ATGAGTTGTTCCAAAGATGATCAAGAGATTCATGAGTTCAAGTTCCATATGAAAGCA 104
QY 21 ThValAAnGlyHtHegluPheGluIleGluGlyGluGlyAArgProTyrgLugly 40
DB 105 AGGTCATATGGCAAGAGTTGAAATTAAGGCAAGGATATATGCAACCCCTGCCGACATACCA 164
QY 41 HisAAnThVallyLeuLyValThThLyGlyGlyProLeuProPheAlaTPAPhile 60
DB 165 CACGTTCGCTAAAGCTTATGATGATCAAGAGTGAAGTTCCTTCATTTGATATAT 224
QY 61 LeuSerProGluPheGluTyrgLySerLyValTyrrVallyshHsPheAlaAPhilePro 80
DB 225 TTGTCAACACAAATTCATGATGAAAGCAAGGATATATGCAACCCCTGCCGACATACCA 284
QY 81 AspTyrtLyshLyLeuSerPhePheProGluGlyPheLystrPgiuAArgValMetAAnPheGlu 100
DB 285 GACATATTAAGTCAATTCCTTCCTGAGGATTTAAATGGAAAGGTCATGAATTTGAA 344
QY 101 AspGlyGlyValValThThValThThGlnApsSerLeuGlnApsGlyCyshPheIleTyrr 120
DB 345 GAGGTCGCGGTGATCTGATATCCCAAGATTCAGTTGAAGACGGCTGTTTCATCTAC 404
QY 121 LyshVallyshPheIleGlyValAAnPheProSerApsGlyProValMetGlnLyshLyThr 140
DB 405 GAGGTCAAGTTCATTTGGGGTGAACCTTCCTTGATGACCTGTTATGACAGAGAGACA 464
QY 141 MetGlyTrpGluAlaSerThThGluArgLeuTyrrProAArgApsGlyValIleuLyGlyGlu 160
```

```
Db 465 CGGGCTGGGAAGCCAGCTCTGAGGCTTGTATCCCGTGAAGGGGTGTGTAAGAGAC 524
Qy 161 ILEHSLVSLALEULYSLYSAAPGLYGLYHSLRYLEUVALGLUPHELYSERILE 180
Db 525 ATCCATATGCTGCTCTGAGGCTGGAAGAGCGGCATTTACCTGTTGAATTCAAAAGTATT 584
Qy 181 TYRMCALALYSLYPR--VALGLNLEUPROGLYTYRTRYTYRVALASPSERYLSLEU 199
Db 585 TACATGCTAAAGAGCCTTCAGTGCAGTTGCCAGGCTTATTATGTTGACTCCAAACTG 644
Qy 200 AAPIETHRSEHISAAGLUASPTYRTHRIEVALGLUGLNTYRGUARGTHRGUGLY 219
Db 645 GATATGACAGGCACACAGAAATTATCACAGTCGTTGACAGATGAAAAAACCCAGGGA 704
Qy 220 ARGHSHISLEUPHELU 225
Db 705 CCGCACCATCCGTTCAATT 722

RESULT 78
US-10-006-076-39
; Sequence 39, Application US/11006076
; Publication No. US20050181506A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
; FILE REFERENCE: 17084-022002/4208
; CURRENT APPLICATION NUMBER: US/11/006,076
; CURRENT FILING DATE: 2004-12-06
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Drosophila species
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (45)...(737)
; OTHER INFORMATION: Nucleotide sequence encoding red fluorescent
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AF272711
; DATABASE ENTRY DATE: 2000-09-26
US-11-006-076-39

Alignment Scores:
Pred. No.: 1,62e-126 Length: 876
Score: 1085.50 Matches: 199
Percent Similarity: 95.584 Conservative: 17
Best Local Similarity: 88.054 Mismatches: 9
Query Match: 89.424 Indels: 1
DB: 10 Gaps: 1

US-10-006-922A-12 (1-225) x US-11-006-076-39 (1-876)
Qy 1 MetArgSerSerLySerValIleLySGluPheMetArgPheLyValArgMetGluGly 20
Db 45 ATGAGTTGTTCCAAGAATGTGATCAAGAGCTTCATGAGTTTCATGAGTAAGA 104
Qy 21 ThrValaenGlyHiaGluPheGluIleGluGlyGluGlyValArgProTyrgluGly 40
Db 105 ACGGTCAATGGGCAAGAGTTTGAATAAAAGCGAAAGCGAGGAGGCTTACGAAGGT 164
```

```
Qy 41 HISAENTHRVALYSLYSLYVALTHRIYSGLYGLYPROLEUPROPHALATRPHEPILE 60
Db 165 CACTGTTCGTAAGACCTTATGTATACAAAGGTGACCTTTGCCATTGCTTTGTATATT 224
Qy 61 LeuSerProGluPheGluIleTyrglySerLyValTyrglyValHiaProAlaAspIlePro 80
Db 225 TTGTACCAACATATTCAATATGAAAGCAAGGTAATGTCAAAACCCGCGCACATACCA 284
Qy 81 AAPIYTRYLYSLYSLYSEPHERPROGLUGLYPHELYSTPGUARGVALMETASNHEGLU 100
Db 285 GACTATTAAGAGCTGCTATTTCTGAGGAAATTTAAATGGAAAGGGTCAATGAACCTTGA 344
Qy 101 AEPGLYGLYVALVALTHRYVALTHRGUNASPSERYLSERLEUGLNAEPGLYCYAPHEILETYR 120
Db 345 GACGCTGCGGTGTTACTGTATCCCAAGATTCAGATTGAAAGAGCGCTTTTCATCTAC 404
Qy 121 LYSVALLYSPHEILEGLYVALASNPHEPROSEAPGLYPROVALMETGLNLYSLYTHR 140
Db 405 GAGGTCAAGTTCAATTGGGGTGAACCTTCTCTGATGACCTGTTATGACAGAGAGACA 464
Qy 141 MetGlyTYRPGUAlaSerThrGluArgLeuTYRProAlaGAPGLYValLeuLYSGLYGLU 160
Db 465 CGGGCTGGGAAGCCAGGCTCTGAGGCTTGTATCCTCGATGAGGGGTCTGAAAGAGAC 524
Qy 161 ILEHSLVSLALEULYSLYSAAPGLYGLYHSLRYLEUVALGLUPHELYSERILE 180
Db 525 ATCCATATGCTGCTCTGAGGCTGGAAGAGCGGCATTTACCTGTTGAATTCAAAAGTATT 584
Qy 181 TYRMCALALYSLYPR--VALGLNLEUPROGLYTYRTRYTYRVALASPSERYLSLEU 199
Db 585 TACATGCTAAAGAGCCTTCAGTGCAGTTGCCAGGCTTATTATGTTGACTCCAAACTG 644
Qy 200 AAPIETHRSEHISAAGLUASPTYRTHRIEVALGLUGLNTYRGUARGTHRGUGLY 219
Db 645 GATATGACAGGCACACAGAAATTATCACAGTCGTTGACAGATGAAAAAACCCAGGGA 704
Qy 220 ARGHSHISLEUPHELU 225
Db 705 CCGCACCATCCGTTCAATT 722

RESULT 79
US-10-724-178-15
; Sequence 15, Application US/10724178
; Publication No. US20040137528A1
; GENERAL INFORMATION:
; APPLICANT: Odyssey Thera, Inc.
; APPLICANT: Michnick, Stephen
; APPLICANT: Macdonald, Marnie
; APPLICANT: Lamerdin, Jane
; TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT
; TITLE OF INVENTION: COMPLEMENTATION ASSAYS
; FILE REFERENCE: ODDY007
; CURRENT APPLICATION NUMBER: US/10/724,178
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US 60/461,133
; PRIOR FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 15
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: "mRFP1", an engineered monomeric form of DsRed FP
; NAME/KEY: CDS
; LOCATION: (1)..(675)
; PUBLICATION INFORMATION:
; AUTHORS: Campbell, R.E. et al
; TITLE: A monomeric red fluorescent protein
; JOURNAL: Proc. Nat'l. Acad. Sci.
; VOLUME: 99
```

```

/ ISSUES: 12
/ PAGES: 7877-82
/ DATE: 2002-06-11
/ RELEVANT RESIDUES: (1) .. (675)
US-10-724-178-15

Alignment Scores:
Pred. No.: 1,566-118 Length: 675
Score: 1021.00 Matches: 192
Percent Similarity: 89.14% Conservative: 5
Best Local Similarity: 86.88% Mismatches: 24
Query Match: 84.10% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-724-178-15 (1-675)

QY 1 MetArgSerSerIySaAnValIleYsGIuPhMeArGpHeLySaValArgMeTGIuGIY 20
DB 1 ATGGCCCTCTCCGAGAGCGTCATCAAGAGATTATCGGCTTCAAGGTGCGCATGAGGGC 60
QY 21 ThrValAsnGIyHIsGluPhHeGIuIleGIuGIyGIuGIyArgProTYrGIuGIY 40
DB 61 TCCGTAAACGGCCACAGATTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 120
QY 41 HisAsnThrValIySeuLySaValIThrLySGLYpProLeuPProHeaLaTPaApIle 60
DB 121 ACCGACACCGCCAAAGCTGAAGGTGACCAAGCGGGCCCCCTGCTTCCGCTGGAGATC 180
QY 61 LeuSerProGInPheGInTYrGIySerLySaValTYrValIyHisProLaApIlePro 80
DB 181 CTGTCCCTCTCCAGAGACGTCATCAAGAGGCTTCAAGTGGAGCGCGTATGAATTCGAG 240
QY 81 AspTYrLySeuLySaSerPheProGIuGIyPheLySTrPGLuArgValMeLaAnPheGIu 100
DB 241 GACTTACTTGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTATGAATTCGAG 300
QY 101 AspGIyGIyValIyValIThrValIThrGInAspSerSerLeuGInAspGIyCySpHeIleTYr 120
DB 301 GACGGCGCGGTGGTGAACCGTGAACCAAGACTCTCCCTGACAGAGCGCGGATTCATCTAC 360
QY 121 LySeuValyPheHeIleGIyValIaAnPheProSerAspGIyProValMeTGIuLySeuThr 140
DB 361 AAGGTGAAGCTGGCGGCGCACCACTTCCCTCCAGCGCCCCCGTAATGCGAAGAAACACC 420
QY 141 MetGIyTrpGIuIaSerThrGIuArgLeuTYrProArgAspGIyValIleuLySGLYGIu 160
DB 421 ATGGGCTGGAGGGCTTCAACCGAGCGATGTACCCCGAGAGCGGGCCCTGAAGGGCGAG 480
QY 161 IleHisLySaIaLeuLySeuLySaAspGIyGIyHIsTYrLeuValGIuPhLySerIle 180
DB 481 ATCAAGATGAGGCTGAAGCTGAAGAGACGGCGGCACATACGACGCGAGTCAAGAACACC 540
QY 181 TyrMeLaIaLySeuLyProValGInLeuProGIyTYrTYrTYrValIaAspSerLySeuAsp 200
DB 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCGGCCCTTACAAAGACCGACATCAAGCTGAC 600
QY 201 IleThrSerHisAsnGIuAspTYrThrIleValGIuGIuInTYrGIuArgThrGIuGIYArg 220
DB 601 ATCACTTCCCAACAGAGACTACACCATGTGTGAACAGTACGAGCGCGCGAGGGCGGC 660
QY 221 His 221
DB 661 CAC 663

```

```

FILE REFERENCE: UC083.1CP2CPI
/ CURRENT APPLICATION NUMBER: US/10/121,258
/ CURRENT FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO: 9
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Dated polypeptide variant
US-10-121-258-9

Alignment Scores:
Pred. No.: 1,576-118 Length: 678
Score: 1021.00 Matches: 192
Percent Similarity: 89.14% Conservative: 5
Best Local Similarity: 86.88% Mismatches: 24
Query Match: 84.10% Indels: 0
DB: 5 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-121-258-9 (1-678)

QY 1 MetArgSerSerIySaAnValIleYsGIuPhMeArGpHeLySaValArgMeTGIuGIY 20
DB 1 ATGGCCCTCTCCGAGAGACGTCATCAAGAGATTATCGGCTTCAAGGTGCGCATGAGGGC 60
QY 21 ThrValAsnGIyHIsGluPhHeGIuIleGIuGIyGIuGIyArgProTYrGIuGIY 40
DB 61 TCCGTAAACGGCCACAGATTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 120
QY 41 HisAsnThrValIySeuLySaValIThrLySGLYpProLeuPProHeaLaTPaApIle 60
DB 121 ACCGACACCGCCAAAGCTGAAGGTGACCAAGCGGGCCCCCTGCTTCCGCTGGAGATC 180
QY 61 LeuSerProGInPheGInTYrGIySerLySaValTYrValIyHisProLaApIlePro 80
DB 181 CTGTCCCTCTCCAGAGACGTCATCAAGAGGCTTCAAGTGGAGCGCGTATGAATTCGAG 240
QY 81 AspTYrLySeuLySaSerPheProGIuGIyPheLySTrPGLuArgValMeLaAnPheGIu 100
DB 241 GACTTACTTGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTATGAATTCGAG 300
QY 101 AspGIyGIyValIyValIThrValIThrGInAspSerSerLeuGInAspGIyCySpHeIleTYr 120
DB 301 GACGGCGCGGTGGTGAACCGTGAACCAAGACTCTCCCTGACAGAGCGGGCCCTGAAGGGCGAG 360
QY 121 LySeuValyPheHeIleGIyValIaAnPheProSerAspGIyProValMeTGIuLySeuThr 140
DB 361 AAGGTGAAGCTGGCGGCGCACCACTTCCCTCCAGCGCCCCCGTAATGCGAAGAAACACC 420
QY 141 MetGIyTrpGIuIaSerThrGIuArgLeuTYrProArgAspGIyValIleuLySGLYGIu 160
DB 421 ATGGGCTGGAGGGCTTCAACCGAGCGATGTACCCCGAGAGCGGGCCCTGAAGGGCGAG 480
QY 161 IleHisLySaIaLeuLySeuLySaAspGIyGIyHIsTYrLeuValGIuPhLySerIle 180
DB 481 ATCAAGATGAGGCTGAAGCTGAAGAGACGGCGGCACATACGACGCGAGTCAAGAACACC 540
QY 181 TyrMeLaIaLySeuLyProValGInLeuProGIyTYrTYrTYrValIaAspSerLySeuAsp 200
DB 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCGGCCCTTACAAAGACCGACATCAAGCTGAC 600
QY 201 IleThrSerHisAsnGIuAspTYrThrIleValGIuGIuInTYrGIuArgThrGIuGIYArg 220
DB 601 ATCACTTCCCAACAGAGACTACACCATGTGTGAACAGTACGAGCGCGCGAGGGCGGC 660
QY 221 His 221

```

```

RESULT 80
US-10-121-258-9
/ Sequence 9, Application US/10121258
/ Publication No. US20030059835A1
/ GENERAL INFORMATION:
/ APPLICANT: Telen, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME

```

Db 661 CAC 663

RESULT 81  
US-10-931-304-9  
/ Sequence 9, Application US/10931304  
/ Publication No. US20050196768A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Teien, Roger  
/ APPLICANT: Campbell, Robert  
/ APPLICANT: Baird, Geoffrey  
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
/ FILE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
/ FILE REFERENCE: 39754-0831CP2CP3  
/ CURRENT APPLICATION NUMBER: US/10/931,304  
/ CURRENT FILING DATE: 2004-08-30  
/ PRIOR APPLICATION NUMBER: 10/209,208  
/ PRIOR FILING DATE: 2002-07-29  
/ PRIOR APPLICATION NUMBER: 10/121,258  
/ PRIOR FILING DATE: 2002-04-10  
/ PRIOR APPLICATION NUMBER: 09/866,538  
/ PRIOR FILING DATE: 2001-05-24  
/ PRIOR APPLICATION NUMBER: 09/794,308  
/ PRIOR FILING DATE: 2001-02-26  
/ NUMBER OF SEQ ID NOS: 110  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 9  
/ LENGTH: 678  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Polynucleotide encoding dimered polypeptide variant  
/ OTHER INFORMATION: "mRFP1"  
US-10-931-304-9

Alignment Scores:  
Pred. No.: 1,57e-118 Length: 678  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88 Mismatches: 24  
Query Match: 84.10% Indels: 0  
Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-9 (1-678)

QY 1 MetAAGSerSerLyAsnValIleLygIuphMeCArgPhelYValArgMetGluGly 20  
DB 1 ATGGGCTCTCCGAGGACGTCATCAAGAGTTCATGCTTCAGAGTCGATGAGGAGC 60  
QY 21 ThrValaAngLyHLeGluPhelGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 TCCGGTAAGCGGCAAGATTCAAGATCGAGGCGAGGGCGGCGCCCTTACGAGGGC 120  
QY 41 HisAanThrValaLyLeuLyValaThrLySgLyGlyProLeuProPhelaATTPAspIle 60  
DB 121 ACCCAGACCGCCAGAGTGAAGTGAACCAAGGGGGCCCTCGCTTGGGAGATC 180  
QY 61 LeuSerProGlnPhelGlnTyrGlySerLyValTyrValaHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCCAGATTCAGATCGGCTCCAGGCGCTTACGTAGACACCCCGCCAGATCCCC 240  
QY 81 AppTyrLyValLeuSerPheProGluGlyPheLyTrgIuArgValaMetAsnPhelGly 100  
DB 241 GATCTACTTGAAGTGTCTTCCCGAGGGCTTCAGTGAAGGCGGTGAATGACTTCGAG 300  
QY 101 AppGlyGlyValaValaThrValaThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGCGGGGTGTGACCTGACCCGAGACTCTCTCCCTCAGAGCGGCGAGTTCATCTAC 360  
QY 121 LySValaLyPheIleGlyValaHisPheProSerAspGlyProValMetGlnLySValaThr 140  
DB 361 AAGGTGAAGCTCGCGGACCAACTTCCCTCCGAGCGGCCCGCTATATGAGAAAGAC 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValaLeuLySgLyGly 160  
DB 421 ATGGGCTGTGGAGGCTTCCAGGAGGATGTAACCCGAGAGCGGCGCCCTGAAGGGCGAG 480  
QY 161 IleHisLySValaLeuLyLeuLyAspGlyGlyHisTyrLeuValaGluPhelYSerIle 180  
DB 481 ATCAAGATGAGCTGTAAGCTGAAGACCGGCGCCACTAGACGCGGAGGTCAAGACACC 540  
QY 181 TyrMetAlaLySValaProValaGluLeuProGluTyrTyrTyrValaAspSerLyAsnApp 200  
DB 541 TACATGGCCAGAACCCGCTGAGCTGCGCGGCTTCAGACCCGAGCATCAAGCTGAGC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValaGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCTCCACACAGAGACTACACCATCTGTGAACGTACGAGCGGCGGAGGGCGGC 660  
QY 221 His 221  
DB 661 CAC 663

RESULT 82  
US-10-931-304-109  
/ Sequence 109, Application US/10931304  
/ Publication No. US20050196768A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Teien, Roger  
/ APPLICANT: Campbell, Robert  
/ APPLICANT: Baird, Geoffrey  
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
/ FILE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
/ FILE REFERENCE: 39754-0831CP2CP3  
/ CURRENT APPLICATION NUMBER: US/10/931,304  
/ CURRENT FILING DATE: 2004-08-30  
/ PRIOR APPLICATION NUMBER: 10/209,208  
/ PRIOR FILING DATE: 2002-07-29  
/ PRIOR APPLICATION NUMBER: 10/121,258  
/ PRIOR FILING DATE: 2002-04-10  
/ PRIOR APPLICATION NUMBER: 09/866,538  
/ PRIOR FILING DATE: 2001-05-24  
/ PRIOR APPLICATION NUMBER: 09/794,308  
/ PRIOR FILING DATE: 2001-02-26  
/ NUMBER OF SEQ ID NOS: 110  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 109  
/ LENGTH: 678  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Polynucleotide encoding dimered polypeptide variant "mGrapel"  
US-10-931-304-109

Alignment Scores:  
Pred. No.: 1.2e-117 Length: 678  
Score: 1014.00 Matches: 191  
Percent Similarity: 89.14% Conservative: 6  
Best Local Similarity: 86.43% Mismatches: 24  
Query Match: 83.53% Indels: 0  
Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-109 (1-678)

QY 1 MetAAGSerSerLyAsnValIleLygIuphMeCArgPhelYValArgMetGluGly 20  
DB 1 ATGGGCTCTCCGAGGACGTCATCAAGAGTTCATGCTTCAGAGTCGATGAGGAGC 60  
QY 21 ThrValaAngLyHLeGluPhelGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 TCCGGTAAGCGGCAAGATTCAAGATCGAGGCGAGGGCGGCGCCCTTACGAGGGC 120  
QY 41 HisAanThrValaLyLeuLyValaThrLySgLyGlyProLeuProPhelaATTPAspIle 60  
DB 121 ACCCAGACCGCCAGAGTGAAGTGAACCAAGGGGGCCCTCGCTTGGGAGATC 180

Query 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIleHisProAlaAspIlePro 80  
DB 181 CTGTCCTCCCTCACTGATGTAACGGCTCAAGGCTCAAGTGAACACCCCTCCACATCCCC 240  
QY 81 AAPPYLYLYLYLeuSerPheProGluGlyPheIysTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTACCTGAAAGCTGCTTCTCCCGAGGCTTCAAGTGGAGCGGTGATGAACTTCGAG 300  
QY 101 AAPPGLYGLYValIleValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGGCGGCTGTGTAACCTGTAACCGAGACTCTCTCCCTGCAAGGACGGGAGTTCACTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140  
DB 361 AAGGTAAAGCTGGCGGCAACAACTTCCCTCCGAGCGCCCGCTATGCAAGAAAGAACCC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleIysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCTCCCGAGCGGCTGTACCCGAGGACGGCGCCCTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180  
DB 481 ATCAAGATGAAGCTGAAGCTGAAGAGACGGGCGCACTACGACCCGAGGCAAGAACACC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCCGTGGAGCTGCGCGCGCTCAACAGCTCAAGCTCAAGCTGAGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAAGAGACTACCATGCTGGAACAGTACGAGCGCGCGAGGCGCGC 660  
QY 221 His 221  
DB 661 CAC 663

RESULT 83  
US-10-931-304-80

/ Sequence 80, Application US/10931304  
/ Publication No. US20050196768A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Tsien, Roger  
/ APPLICANT: Campbell, Robert  
/ APPLICANT: Baird, Geoffrey  
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
/ FILE REFERENCE: 39754-0831CP2CP3  
/ CURRENT APPLICATION NUMBER: US/10/931.304  
/ PRIOR FILING DATE: 2004-08-30  
/ PRIOR APPLICATION NUMBER: 10/209,208  
/ PRIOR FILING DATE: 2002-07-29  
/ PRIOR APPLICATION NUMBER: 10/121,258  
/ PRIOR FILING DATE: 2002-04-10  
/ PRIOR APPLICATION NUMBER: 09/866,538  
/ PRIOR FILING DATE: 2001-05-24  
/ PRIOR APPLICATION NUMBER: 09/794,308  
/ PRIOR FILING DATE: 2001-02-26  
/ NUMBER OF SEQ ID NOS: 110  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 80  
/ LENGTH: 678  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
/ OTHER INFORMATION: "mrFp1.1"  
US-10-931-304-80

Alignment Scores:  
Pred. No.: 2,156-117 Length: 678  
Score: 1012.00 Matches: 190  
Percent Similarity: 88.69% Conservative: 6  
Best Local Similarity: 85.97% Mismatches: 25

Query Match: 83.36% Indels: 0  
DB: 9 Gaps: 0  
US-10-006-922a-12 (1-225) x US-10-931-304-80 (1-678)

QY 1 MetArgSerSerIysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGGCTCTCCGAGAGAGTATCAAGAGGTTCAAGCCCTTCAAGGTGCGATGAGGGCC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluArgProTyrGluGly 40  
DB 61 TCCGTGAACGGCCACAGATTCAAGATCAAGGCGAGGCGAGAGGCGCGCCCTTACAGAGGCG 120  
QY 41 HisAsnThrValIysLeuLysValThrIysGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 ACCGAGACCGCAAGCTGAAGGTGAACAAAGGCGGCGCCCTTCCCTCCGAGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIleHisProAlaAspIlePro 80  
DB 181 CTGTCCTCCCTCACTGATGTAACGGCTCAAGGCTCAAGTGAACACCCCTCCACATCCCC 240  
QY 81 AAPPYLYLYLYLeuSerPheProGluGlyPheIysTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTACCTGAAAGCTGCTTCTCCCGAGGCTTCAAGTGGAGCGGTGATGAACTTCGAG 300  
QY 101 AAPPGLYGLYValIleValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGGCGGCTGTGTAACCTGTAACCGAGACTCTCTCCCTGCAAGGACGGGAGTTCACTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140  
DB 361 AAGGTAAAGCTGGCGGCAACAACTTCCCTCCGAGGCGCCCGCTATGCAAGAAAGAACCC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleIysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCTCCCGAGCGGATGTACCCGAGGACGGCGCCCTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180  
DB 481 ATCAAGATGAAGCTGAAGCTGAAGAGACGGGCGCACTACGACCCGAGGCAAGAACACC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCCGTGGAGCTGCGCGCGCTCAACAGCTCAAGCTCAAGCTGAGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAAGAGACTACCATGCTGGAACAGTACGAGCGCGCGAGGCGCGC 660  
QY 221 His 221  
DB 661 CAC 663

## RESULT 84

US-10-931-304-105  
/ Sequence 105, Application US/10931304  
/ Publication No. US20050196768A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Tsien, Roger  
/ APPLICANT: Campbell, Robert  
/ APPLICANT: Baird, Geoffrey  
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
/ FILE REFERENCE: 39754-0831CP2CP3  
/ CURRENT APPLICATION NUMBER: US/10/931.304  
/ PRIOR FILING DATE: 2004-08-30  
/ PRIOR APPLICATION NUMBER: 10/209,208  
/ PRIOR FILING DATE: 2002-07-29  
/ PRIOR APPLICATION NUMBER: 10/121,258  
/ PRIOR FILING DATE: 2002-04-10  
/ PRIOR APPLICATION NUMBER: 09/866,538  
/ PRIOR FILING DATE: 2001-05-24  
/ PRIOR APPLICATION NUMBER: 09/794,308

```
/ PRIOR FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 105
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Dated polypeptide variant "mHoneydew"
US-10-931-304-105

Alignment Scores:
Pred. No.: 9,2e-117 Length: 678
Score: 1007.00 Matches: 188
Percent Similarity: 88.69% Conservative: 8
Best Local Similarity: 85.07% Mismatches: 25
Query Match: 82.95% Indels: 0
DB: 9 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-105 (1-678)

QY 1 MetAAGSerSerLYaAsnValIlelysgLuphMeAArgPhelyValAArgMetGluGly 20
DB 1 ATGGGCTCTCCGAGGAGCTGATCAAGAGGTTCAAGCTTCAAGGTGCGATGAGGAGGC 60
QY 21 ThrValAaenGLYhIeGLuphneGLuIleGLuGLuGLuGLuGLuGLuGLuGLuGLuGLu 40
DB 61 TCCGTGAACGGCCACAGATTCCAGATCGAGGCGAGGCGGAGGCGGCGGCGGCGGCGGCG 120
QY 41 HisAenThrValIlelysgLuphneGLuIleGLuGLuGLuGLuGLuGLuGLuGLuGLuGLu 60
DB 121 ACCGAGACCGCCAGAGCTGAAGGTGACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCG 180
QY 61 LeuSerProGLInPhneGLInTyGLySerIySvalTyValIyHisProLAAspIlePro 80
DB 181 CTGTCCCTCCAGATTGATGCGGAGCTTCAAGGCTTCAAGTAAACACCCGCGACATCCCC 240
QY 81 AepTYrLYsLYsLeuSerPheProGLuGLyPheLYsTPRGLAArgValMeAAsnPhneGLu 100
DB 241 GACTACTTGAAAGCTGTCTTCCCGAGGCGTTCAAGTGGAGCGGCGTGAATTTCCAG 300
QY 101 AepGLyGLyValIleValIleThrValIleThrIleAepSerSerLeuGLInAepGLyCyAspHeIleTYr 120
DB 301 GACGGCGGGGTGGTACCGTGACCCAGGACTCTCCCTGACGAGCGGCGAGTTCAATCTAC 360
QY 121 LYsValIySPheIleGLyValAAsnPheProSerAAspGLyProValMeGLInIySvalTYr 140
DB 361 AAGGTGAAGCTGGCGGACCAATTCCCTCCGACGCGCGCGTAAATGCAAGAAAGACC 420
QY 141 MetGLyTPRGLuAAserThrGLuArgLeuTYrProAArgAAspGLyValLeuIySGLyGLu 160
DB 421 ATGGGCTGGGCGGCGCCACCAAGCGGATGTACCCCGAGAGCGGCGGCGGCGGCGGCGG 480
QY 161 IleHisLYsValIleuLYsLeuLYsAAspGLyGLyHisTYrLeuValIleGluPheLYsSerIle 180
DB 481 ATCAAGATAGAGCTGAAGCTGAAGAGCGGCGGCGCTTCAAGCGGCGGCGGCGGCGGCGG 540
QY 181 TYrMeAAlaLYsLYsProValIleGluPheProGLyTYrTYrTYrValAAspSerIySLeuAAsp 200
DB 541 TACATGGCCCAAGAGCGCGTGCAGCGCGCGCGCGCTTCAAGATTGACGCGAAGCGTGCAG 600
QY 201 IleThrSerHisAAsnGLuAAspTYrThrIleValIleGluGLInTyGLuArgThrGLuGLu 220
DB 601 ATCACTCTCCCAACAGAGACTACACCATCTGTGAACAGTACAGAGCGGCGGCGGCGGCGG 660
QY 221 His 221
DB 661 CAC 663

RESULT 85
US-10-931-304-101
/ Sequence 101, Application US/10931304
/ Publication No. US20050196768A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Tsien, Roger
/ APPLICANT: Campbell, Robert
/ APPLICANT: Baird, Geoffrey
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: 39754-083ICP2CP3
/ CURRENT APPLICATION NUMBER: US/10/931,304
/ PRIOR FILING DATE: 2004-08-30
/ PRIOR APPLICATION NUMBER: 10/209,208
/ PRIOR FILING DATE: 2002-07-29
/ PRIOR APPLICATION NUMBER: 10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 101
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Dated polypeptide variant "mTangerine"
US-10-931-304-101

Alignment Scores:
Pred. No.: 3,94e-116 Length: 678
Score: 1002.00 Matches: 189
Percent Similarity: 88.24% Conservative: 6
Best Local Similarity: 85.52% Mismatches: 26
Query Match: 82.54% Indels: 0
DB: 9 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-101 (1-678)
```

```
QY 1 MetAAGSerSerLYaAsnValIlelysgLuphMeAArgPhelyValAArgMetGluGly 20
DB 1 ATGGGCTCTCCGAGGAGCTGATCAAGAGGTTCAAGCTTCAAGGTGCGATGAGGAGGC 60
QY 21 ThrValAaenGLYhIeGLuphneGLuIleGLuGLuGLuGLuGLuGLuGLuGLuGLuGLu 40
DB 61 TCCGTGAACGGCCACAGATTCCAGATCGAGGCGAGGCGGCGGCGGCGGCGGCGGCGGCG 120
QY 41 HisAenThrValIlelysgLuphneGLuIleGLuGLuGLuGLuGLuGLuGLuGLuGLuGLu 60
DB 121 ACCGAGACCGCCAGAGCTGAAGGTGACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCG 180
QY 61 LeuSerProGLInPhneGLInTyGLySerIySvalTyValIyHisProLAAspIlePro 80
DB 181 CTGTCCCTCCAGATTGATGCGGAGCTTCAAGGCTTCAAGTAAACACCCGCGACATCCCC 240
QY 81 AepTYrLYsLYsLeuSerPheProGLuGLyPheLYsTPRGLuAArgValMeAAsnPhneGLu 100
DB 241 GACTACTTGAAAGCTGTCTTCCCGAGGCGTTCAAGTGGAGCGGCGGCGTGAATTTCCAG 300
QY 101 AepGLyGLyValIleValIleThrValIleThrIleAepSerSerLeuGLInAepGLyCyAspHeIleTYr 120
DB 301 GACGGCGGGGTGGTACCGTGACCCAGGACTCTCCCTGACGAGCGGCGGCGGCGGCGGCGG 360
QY 121 LYsValIySPheIleGLyValAAsnPheProSerAAspGLyProValMeGLInIySvalTYr 140
DB 361 AAGGTGAAGCTGGCGGACCAATTCCCTCCGACGCGCGCGTAAATGCAAGAAAGACC 420
QY 141 MetGLyTPRGLuAAserThrGLuArgLeuTYrProAArgAAspGLyValLeuIySGLyGLu 160
DB 421 ATGGGCTGGGCGGCGCCACCAAGCGGATGTACCCCGAGAGCGGCGGCGGCGGCGGCGG 480
QY 161 IleHisLYsValIleuLYsLeuLYsAAspGLyGLyHisTYrLeuValIleGluPheLYsSerIle 180
DB 481 ATCAAGATAGAGCTGAAGCTGAAGAGCGGCGGCGCTTCAAGCGGCGGCGGCGGCGGCGG 540
```

```
QY 181 Tymealalalyeprovalglneuproglytyrttyrtvalaserylseuap 200
DB 541 TACATGCGCAAGAGCCCTGCAAGCTGCGCGGCGCTTACAAAGCCACATCAAGCTGAGC 600
QY 201 Iietrsehisaaagluasaptyrtthriilevalgluglntyrgluarthrngluglyaxg 220
DB 601 ATCACCCTCCCAACAGAGGACTACACCATGTGAAATTGTGAGAGCGCCGAGAGGCGCG 660
QY 221 His 221
DB 661 CAC 663

RESULT 86
US-10-931-304-86
; Sequence 86, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding Dated polypeptide variant "Ors4-9"
US-10-931-304-86

Alignment Scores:
Pred. No.: 4,3e-115 Length: 711
Score: 994.00 Matches: 186
Percent Similarity: 86.84% Conservative: 5
Best Local Similarity: 86.51% Mismatches: 24
Query Match: 81.88% Indels: 0
DB: 9 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-86 (1-711)
QY 7 Valilleysgluphemetaargpnehyvalargmetgluglythrvalaenglyhisglu 26
DB 34 ATCATCAAGAGAGTTCATGCGCTTCAAGGTACGATGAGGGCTCCGTGAACGGCCAGAG 93
QY 27 Phegluilegluglntyrgluglyaxgprotyrttyrttyrtthriilevalaenglyhisglu 46
DB 94 TTCGAGATTCAGAGGCGGAGGCGGAGGCGCGCTTCAAGAGGCGCCAGACCGCCAAAGTTG 153
QY 47 lyevalthrlysglyglpyleuprophealatrphapriileuseerproglupheglu 66
DB 154 AAGGTGACCAAGGTTGCGCCCTTCCCTTCCGCGGACATCCGTCCCTCAGTTCAACC 213
QY 67 tyrttyrttyrttyrttyrttyrttyrttyrttyrttyrttyrttyrttyrttyrttyrt 86
DB 214 TACGGCTCAAGGCGCTACGTGAAGACACCGCCGACATCCCGACTACTTGAAAGCTGACC 273
QY 87 Pheprogluglypnehygluglyaxgvalmetanphegluaprglyglvalvalthr 106
DB 274 TTCGCCAGAGGCTTCAAGTGGGAACGGGTGATGAATTCGAGAGCGGCGGTGGTGAACC 333
```

```
QY 107 ValthrgluapserSerleuGlunapgllyCyPheilleTyrtlyeVallyasPheillegly 126
DB 334 GTGACACAGAGACTCTCTCTGCAAGACGGCGAGTTCACTTCAAGTGAAGCTGGCGGCG 393
QY 127 ValaapneProSerAspGlyProvalmetgluylslyethmetglYTrgluAlaser 146
DB 394 ACCAACTTCCCTCCGAGCGGCCCGGTAATGCAAGAAAGACCATGGGCTGGAGGCTTCC 453
QY 147 thrgluabgleutyrtProArGAspGlyValleuylsglyglunillehislyalaleuyls 166
DB 454 TCCGAGCGGATGTACCCCGCAGAGCGCGCTGAAGGGGAGATCAAGATGAGGCTGAAG 513
QY 167 leuylasapgllygllysttyrtleuvalglupheylseriletyrmetalalyserPro 186
DB 514 CTGAAGAGAGCGGCGCATCAAGCGCTGAGTCMAAGACCACTTCAAGGCCAAGAGCC 573
QY 187 ValgluLeuproglytyrttyrtvalaserylseuapriilethrserhisaaenglu 206
DB 574 GTGAGCTCCCGGCGCTTCAAGAGTCAAGATCAAGCTGAGATCACTCCCAACAGAG 633
QY 207 Aspyrtthriilevalgluglntyrgluatgthrngluglyaxghis 221
DB 634 GACTACACCATGCTGGAGACGTAAGAACCGCCGAGGCGCGCAC 678

RESULT 87
US-10-931-304-90
; Sequence 90, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 90
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynuc. encoding polypeptide variant "mFRFP (P206) (mGrape2)"
US-10-931-304-90

Alignment Scores:
Pred. No.: 7,69e-115 Length: 711
Score: 992.00 Matches: 185
Percent Similarity: 88.84% Conservative: 6
Best Local Similarity: 86.05% Mismatches: 24
Query Match: 81.71% Indels: 0
DB: 9 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-90 (1-711)
QY 7 Valilleysgluphemetaargpnehyvalargmetgluglythrvalaenglyhisglu 26
DB 34 GTATCAAGAGAGTTCATGCGCTTCAAGGTACGATGAGGGCTCCGTGAACGGCCAGAG 93
QY 27 Phegluilegluglntyrgluglyaxgprotyrttyrttyrtthriilevalaenglyhisglu 46
DB 94 TTCGAGATTCAGAGGCGGAGGCGGAGGCGCGCTTCAAGAGGCGCCAGACCGCCAAAGTTG 153
```

```

QY 47 LysValThrIleGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66
DB 154 AAGGAGACCAAGGGGGCCCTGGCTTGGCTGGAGATCATCTGCTCCCTCAGTTCAAG 213
QY 67 TyrGlySerIleValTyrValIleHisProAlaAspIleProAspTyrIleValIleSer 86
DB 214 TACGGCTCCAAAGGCTTCAAGTGAAGCAGCCCTGACATCCCGACATCAATGAAGCTGTCC 273
QY 87 PheProGlnGlyPheIleTrrGluArgValMetAsnPheGlnAspGlyValIleValIleThr 106
DB 274 TTCCTCCAGAGGCTTCAAGTGAAGCAGCTGATGAATCTTGAGAGACGGCGCTGGTGAACC 333
QY 107 ValTrnGlnAspSerSerLeuGlnAspGlyCysPheIleTyrIleValIleHisPheIleGly 126
DB 334 GAGACCCCAAGAGATCTCTCCCTGGAGAGACGGCGAGTTATCTACAAAGGTGAAGCTGACGGC 393
QY 127 ValAsnPheProSerAspGlyProValMetGlnIleValIleThrMetGlyTrrGluAlaSer 146
DB 394 ACCAATCTCCCTCCAGACGGCGCCCGTAAATGCAAGAAAGACATGGGCTGGAGGCTTCC 453
QY 147 ThrGluArgLeuTyrProArgAspGlyValLeuIleGlyGluIleHisValIleLeuIle 166
DB 454 TCCGAGCGGCTTGAACCCCAAGAGACGGCGCTGAAGGGCGAGGTCAAGATGAAGCTGAAG 513
QY 167 LeuIleAspGlyGlyHisTyrIleValGluPheIleTyrMetAlaIleValIlePro 186
DB 514 CTGAAGGAGAGGGGGCCACTACAGACGGCGAGGTCAAGACCACTACATGGCCAAAGAGCC 573
QY 187 ValGlnLeuProGlyTyrTyrTyrValAlaSerIleValLeuAspIleThrSerHisAnglu 206
DB 574 GTGCAGCTGCGCGCGCTCAAGACTCGACTCAAGCTGACATCACTCCCAACAGAG 633
QY 207 AspTyrThrIleValGluGlnTyrGluArgTrnGluIleArgHis 221
DB 634 GACTACACCATGTGGAAAGTACGAGCGCGCCGAGGGCGCCAC 678

```

```

RESULT 88
US-10-931-304-103
; Sequence 103, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mOrange (mOP1
US-10-931-304-103

```

```

Alignment Scores:
Pred. No.: 1,84e-114 Length: 711
Score: 989.00 Matches: 185
Percent Similarity: 88.37% Conservative: 5
Best Local Similarity: 86.05% Mismatches: 25
Query Match: 81.47% Indels: 0

```

```

DB: 9 Gaps: 0
US-10-006-922A-12 (1-225) x US-10-931-304-103 (1-711)
QY 7 ValIleIleGluPheMetArgPheIleValArgMetGluGlyThrValIleGlyHisGlu 26
DB 34 ATCAATCAAGAGATTATGCTGCTTCAAGGTGCGCAATGAGAGGGCTCGGTGAACGGCCACGAG 93
QY 27 PheGluIleGluGluGluGluGluValArgProTyrGluIleHisAsnThrValIleLeu 46
DB 94 TTCGATGAGAGGGAGAGGCGAGGGCGCCCTTACAGAGGCTTTCAGACCCCTTAAGCTTG 153
QY 47 LysValThrIleGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66
DB 154 AAGGAGACCAAGGGGGCCCTGGCTTGGCTGGAGATCATCTGCTCCCTCAGTTCAAG 213
QY 67 TyrGlySerIleValTyrValIleHisProAlaAspIleProAspTyrIleValIleSer 86
DB 214 TACGGCTCCAAAGGCTTCAAGTGAAGCAGCCCTGACATCCCGACATCAATGAAGCTGTCC 273
QY 87 PheProGlnGlyPheIleTrrGluArgValMetAsnPheGlnAspGlyValIleValIleThr 106
DB 274 TTCCTCCAGAGGCTTCAAGTGAAGCAGCTGATGAATCTTGAGAGACGGCGCTGGTGAACC 333
QY 107 ValTrnGlnAspSerSerLeuGlnAspGlyCysPheIleTyrIleValIleHisPheIleGly 126
DB 334 GAGACCCCAAGAGATCTCTCCCTGGAGAGACGGCGAGTTATCTACAAAGGTGAAGCTGACGGC 393
QY 127 ValAsnPheProSerAspGlyProValMetGlnIleValIleThrMetGlyTrrGluAlaSer 146
DB 394 ACCAATCTCCCTCCAGACGGCGCCCGTAAATGCAAGAAAGACATGGGCTGGAGGCTTCC 453
QY 147 ThrGluArgLeuTyrProArgAspGlyValLeuIleGlyGluIleHisValIleLeuIle 166
DB 454 TCCGAGCGGCTTGAACCCCAAGAGACGGCGCTGAAGGGCGAGATCAAGATGAAGCTGAAG 513
QY 167 LeuIleAspGlyGlyHisTyrIleValGluPheIleTyrMetAlaIleValIlePro 186
DB 514 CTGAAGGAGAGGGGGCCACTACAGACGGCGAGGTCAAGACCACTACAAAGGCCAAAGAGCC 573
QY 187 ValGlnLeuProGlyTyrTyrTyrValAlaSerIleValLeuAspIleThrSerHisAnglu 206
DB 574 GTGCAGCTGCGCGCGCTCAAGACTCGACTCAAGCTGACATCACTCCCAACAGAG 633
QY 207 AspTyrThrIleValGluGlnTyrGluArgTrnGluIleArgHis 221
DB 634 GACTACACCATGTGGAAAGTACGAGCGCGCCGAGGGCGCCAC 678

```

```

RESULT 89
US-10-931-304-84
; Sequence 84, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 711

```



```
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Derived polypeptide variant "mRPL5"
US-10-931-304-84

Alignment Scores:
Pred. No.: 2,466-114 Length: 711
Score: 988.00 Matches: 184
Percent Similarity: 88.374 Conservative: 6
Best Local Similarity: 85.584 Mismatches: 25
Query Match: 81.384 Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-84 (1-711)

QY 7 ValIleLeuGluPheMetArgPheLeuValArgMetGluGlyThrValAlaGlyHisGlu 26
   :::::
DB 34 ATCATCAAGAGATTCAATGCGCTTCAGAGGTCACATGAGAGGCTCCGTGAACGGCCACGAG 93

QY 27 PheGluIleGluGlyGluGlyGluGlyArgProTyrGlyGlyHisAsnThrValIleu 46
   :::::
DB 94 TTCGAAATCGAAGGCGCAGGCGGCGGCGGCGCTTACAGAGGCGACCCAGACCGCCAGCGTG 153

QY 47 LysValThrIleGlyGlyProLeuProPheAlaTyrAspIleLeuSerProGlnPheGln 66
   :::::
DB 154 AAGGTACCAAGGGGTGGCCCTGCGCTTCGCTGGGACATCTCTGCTCCCTCAGTTTCATG 213

QY 67 TyrGlySerLeuValIleValIleValIleValIleValIleValIleValIleValIle 86
   :::::
DB 214 TACGGCTCCAGAGGCTTCAAGTGGAGCGCGTGAATGAATCTTCAGAGGCGCGCGTGAC 273

QY 87 PheProGluGlyPheLeuTyrGluArgValMetAsnPheGluAspGlyValIleThr 106
   :::::
DB 274 TTCCTCCAGAGGCTTCAAGTGGAGCGCGTGAATGAATCTTCAGAGGCGCGCGTGAC 333

QY 107 ValThrGlnAspSerLeuGlnAspGlyCysPheIleTyrIleValIleValIleGly 126
   :::::
DB 334 GTGACCAAGAGCTCTCCCTGCGAGGAGCGCGAGTTCATCTACAGAGTGAAGTGGCGGCG 393

QY 127 ValAsnPheProSerAspGlyProValMetGlnIleValIleValIleValIleValIle 146
   :::::
DB 394 ACCAACTTCCCTCCGAGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 453

QY 147 ThrGluArgLeuTyrProArgAspGlyValIleuValIleValIleHisAlaLeuLys 166
   :::::
DB 454 TCCGAGGAGATGATACCCCGAGAGCGGCGGCGCTTGAAGGCGAGATCAAGCAGAGCTGAG 513

QY 167 LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186
   :::::
DB 514 CTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 573

QY 187 ValGlnLeuProGlyTyrTyrValIleAspSerIleLeuAspIleThrSerHisAsnGlu 206
   :::::
DB 574 GTCAAGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 633

QY 207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221
   :::::
DB 634 GACTACACATCGTGAACAGTACGAACGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCG 678

RESULT 90
US-10-931-304-93
/ Sequence 93. Application US/10931304
/ Publication No. US20050196768A1
/ GENERAL INFORMATION:
/ APPLICANT: Telen, Roger
/ APPLICANT: Campbell, Robert
/ APPLICANT: Baird, Geoffrey
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: 39754-0831C2C93
/ CURRENT APPLICATION NUMBER: US/10/931.304
/ CURRENT FILING DATE: 2004-08-30
```

```
/ PRIOR APPLICATION NUMBER: 10/209,208
/ PRIOR FILING DATE: 2002-07-29
/ PRIOR APPLICATION NUMBER: 10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 93
/ LENGTH: 711
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Derived polypeptide variant "mRPL2 (mCherry)"
US-10-931-304-93

Alignment Scores:
Pred. No.: 3,296-114 Length: 711
Score: 987.00 Matches: 184
Percent Similarity: 88.844 Conservative: 7
Best Local Similarity: 85.584 Mismatches: 24
Query Match: 81.304 Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-93 (1-711)

QY 7 ValIleLeuGluPheMetArgPheLeuValArgMetGluGlyThrValAlaGlyHisGlu 26
   :::::
DB 34 ATCATCAAGAGATTCAATGCGCTTCAAGAGGTCACATGAGAGGCTCCGTGAACGGCCACGAG 93

QY 27 PheGluIleGluGlyGluGlyGluGlyArgProTyrGlyGlyHisAsnThrValIleu 46
   :::::
DB 94 TTCGAAATCGAAGGCGCAGGCGGCGGCGGCGGCGCTTACAGAGGCGACCCAGACCGCCAGCGTG 153

QY 47 LysValThrIleGlyGlyProLeuProPheAlaTyrAspIleLeuSerProGlnPheGln 66
   :::::
DB 154 AAGGTACCAAGGGGTGGCCCTGCGCTTCGCTGGGACATCTCTGCTCCCTCAGTTTCATG 213

QY 67 TyrGlySerLeuValIleValIleValIleValIleValIleValIleValIleValIle 86
   :::::
DB 214 TACGGCTCCAGAGGCTTCAAGTGGAGCGCGTGAATGAATCTTCAGAGGCGCGCGTGAC 273

QY 87 PheProGluGlyPheLeuTyrGluArgValMetAsnPheGluAspGlyValIleThr 106
   :::::
DB 274 TTCCTCCAGAGGCTTCAAGTGGAGCGCGTGAATGAATCTTCAGAGGCGCGCGTGAC 333

QY 107 ValThrGlnAspSerLeuGlnAspGlyCysPheIleTyrIleValIleValIleGly 126
   :::::
DB 334 GTGACCAAGAGCTCTCCCTGCGAGGAGCGCGAGTTCATCTACAGAGTGAAGTGGCGGCGG 393

QY 127 ValAsnPheProSerAspGlyProValMetGlnIleValIleValIleValIleValIle 146
   :::::
DB 394 ACCAACTTCCCTCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 453

QY 147 ThrGluArgLeuTyrProArgAspGlyValIleuValIleValIleHisAlaLeuLys 166
   :::::
DB 454 TCCGAGGAGATGATACCCCGAGAGCGGCGGCGCTTGAAGGCGAGATCAAGCAGAGCTGAG 513

QY 167 LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186
   :::::
DB 514 CTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 573

QY 187 ValGlnLeuProGlyTyrTyrValIleAspSerIleLeuAspIleThrSerHisAsnGlu 206
   :::::
DB 574 GTCAAGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 633

QY 207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221
   :::::
DB 634 GACTACACATCGTGAACAGTACGAACGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 678

RESULT 91
```

```
US-10-931-304-95
; Sequence 95, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding Dered polypeptide variant "mYOFp (74-11)"
; NAME/KEY: misc_feature
; LOCATION: 185
; OTHER INFORMATION: n = A,T,C or G
US-10-931-304-95

Alignment Scores:
Pred. No.: 4,37e-114 Length: 708
Score: 986.00 Matches: 184
Percent Similarity: 87.91% Conservative: 5
Best Local Similarity: 85.58% Mismatches: 26
Query Match: 81.22% Indels: 0
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-95 (1-708)
QY 7 ValIlelyGluPhMeArqPhelyValAArgMetGluGlyThrValAanglyHsGlu 26
   ::::::::::::::::::::
DB 34 ATCATCAAGAGATTATGCGCTTCAAGGTGCGCATGAGGGCTCGGTGAACGGCCACGAG 93
QY 27 PheGluIleGluGlyGluGlyAArgProTyArgGlyHsAAsnThrValIysIeu 46
   ::::::::::::::::::::
DB 94 TTCGAGATCGAGGGCGAGGGCCAGGGCCCTTACGAGGGCTTTCAGACCGCTTAAGCTG 153
QY 47 LysValThrIysGlyGlyProIeuProPheAlaTrpAspIleIeuSerProGlnPheGln 66
   ::::::::::::::::::::
DB 154 AAGGAGACCAAGGGTGGCCCTGCGCTTCGACTGGGACATCTGCTCCCTCAAGTTCAAC 213
QY 67 TyrGlySerIysValIyValIysHsIProAlaAspIleProAspTyTrIysIysIeuSer 86
   ::::::::::::::::::::
DB 214 TACGGCTCCAAAGGCTCACTGAAGCAACCCCGCGCATCCCGGACTTGAAGCTGTCC 273
QY 87 PheProGluGlyPheIysTrpGluArgValIleAAsnPheGluAAspGlyValIyValIThr 106
   ::::::::::::::::::::
DB 274 TTCCTCCGAGGGCTTCAAGTGGAGCGCGTGAATCTTCGAGAGCGGGCGGCTGGTGAAC 333
QY 107 ValThrGlnAAspSerSerIeuGlnAAspGlyCyPheIleTyTrIysValIysPheIleGly 126
   ::::::::::::::::::::
DB 334 GTCGACCCAGAGATCTCTCTCTGCAAGACGGCGGACTTATTAAGGTGAAGCTGGCGGCG 393
QY 127 ValAAsnPheProSerAAspGlyProValIleMetGlnIyValIysThrMetGlyTrpGluAAser 146
   ::::::::::::::::::::
DB 394 ACCAAGTTCCTCCCTCCGAGCGGGCCCGTAATGCAAGAAAGACATAGGGGTGGAGGGCTTCC 453
QY 147 ThrGluArgIeuTyProArgAAspGlyValIleuIyGlyIleHsIlyValIleuIyS 166
```

```
DB 454 TCCGAGCGGATGTACCCCGAGAGACGGCCCTGAAAGGCGAGATCAAGATGAGCGTGAAG 513
   ::::::::::::::::::::
QY 167 LeuIysAAspGlyGlyHsIyTrIeuValIuPhelySerIleTyTrMetAlaIyIysPro 186
   ::::::::::::::::::::
DB 514 CTGAAGGACGGCGGCTCACTACCACTCCAGGTCAAGACCACTTACAGGCGCAAGAGCCC 573
QY 187 ValGlnIeuProGlyTyTrTyTrValIAspSerIysIeuAspIleThrSerHsAanglu 206
   ::::::::::::::::::::
DB 574 GTCGAGCTGCCCCGGGCTTACATCTGCGCATCAAGTTGACATACCTTCCCAACAGAG 633
QY 207 AspTyThrIleValIuGluGlnTyArgIuArgThrGluIyArgHs 221
   ::::::::::::::::::::
DB 634 GACTACACCATGTGTGAACAGTACGACGAGCGGCGGAGCGGCGGCAC 678

RESULT 92
US-10-931-304-97
; Sequence 97, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding Dered polypeptide variant "mKOFp (A2/6-6)"
US-10-931-304-97

Alignment Scores:
Pred. No.: 1,93e-112 Length: 711
Score: 973.00 Matches: 182
Percent Similarity: 87.44% Conservative: 6
Best Local Similarity: 84.65% Mismatches: 27
Query Match: 80.15% Indels: 0
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-97 (1-711)
QY 7 ValIlelyGluPhMeArqPhelyValAArgMetGluGlyThrValAanglyHsGlu 26
   ::::::::::::::::::::
DB 34 ATCATCAAGAGATTATGCGCTTCAAGGTGCGCATGAGGGCTCGGTGAACGGCCACGAG 93
QY 27 PheGluIleGluGlyGluGlyAArgProTyArgGlyHsAAsnThrValIysIeu 46
   ::::::::::::::::::::
DB 94 TTCGAGATCGAGGGCGAGGGCCAGGGCCCTTACGAGGGCAACCAAGCGCCAAAGCTG 153
QY 47 LysValThrIysGlyGlyProIeuProPheAlaTrpAspIleIeuSerProGlnPheGln 66
   ::::::::::::::::::::
DB 154 AAGGAGACCAAGGGTGGCCCTGCGCTTCGACTGGGACATCTTAAACCCCAACTTCAAC 213
QY 67 TyrGlySerIysValIyValIysHsIProAlaAspIleProAspTyTrIysIysIeuSer 86
   ::::::::::::::::::::
DB 214 TACGGCTCCAAAGGCTCACTGAAGCAACCCCGCGCATCCCGGACTTGAAGCTGTCC 273
QY 274 PheProGluGlyPheIysTrpGluArgValIleAAsnPheGluAAspGlyValIyValIThr 106
```

DB 274 TTCCTCCGAGGGCTTCAAGTGGAGCGCGCTGATGAATCTTGAGAGCGCGCGCTGGTACC 333  
QY 107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLeuValIleAspHleGly 126  
DB 334 GTGACCCGAGACTCTCTCCCTGCGAGCGCGGAGTTCACTTCAAGGTAGTGAAGCTGGCGGC 393  
QY 127 ValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTyrGlnLysSer 146  
DB 394 ACCAATCTCCCTCCGAGCGCGCGCTGATGAGAGAGAACATCGGCTGGAGGCTTCC 453  
QY 147 ThrGlnAspLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166  
DB 454 TCCGAGCGGATGATACCCCGAGAGCGCGCGCTGAGAGGAGATCAAGTGAAGCTGAAG 513  
QY 167 LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186  
DB 514 CTGAAGAGAGCGGCGCATACGACGCTGAGGTCAAGACCACTTCAAGGCCAAGAGCCC 573  
QY 187 ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206  
DB 574 GTGCGACTCCCGCGCGCTCAACATCGTCGGCATCAAGTTGGAATCACTCCCAACAGAG 633  
QY 207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221  
DB 634 GACTACACCATCGTGAACGTAGAACGCGCGCGCGCCAC 678  
RESULT 93  
US-10-931-304-99  
Sequence 99, Application US/10931304  
Publication No. US20050196768A1  
GENERAL INFORMATION:  
APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert  
APPLICANT: Baird, Geoffrey  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
FILE REFERENCE: 39754-0831CP2CP3  
CURRENT APPLICATION NUMBER: US/10/931,304  
CURRENT FILING DATE: 2004-08-30  
PRIOR APPLICATION NUMBER: 10/209,208  
PRIOR FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 99  
LENGTH: 711  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Polynucleotide encoding Dared polypeptide variant "mstrawberry"  
US-10-931-304-99  
Alignment Scores:  
Pred. No.: 1,93e-112 Length: 711  
Score: 973.00 Matches: 182  
Percent Similarity: 87.44% Conservative: 6  
Best Local Similarity: 84.65% Mismatches: 27  
Query Match: 80.15% Indels: 0  
Gaps: 0  
DB: 9  
US-10-006-922a-12 (1-225) x US-10-931-304-99 (1-711)  
QY 7 ValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGlyHisGlu 26  
DB 34 ATCATCAAGAGATTCAATGCGCTTCAAGGTGCGCATGAGAGGCTCGTGAAGGCCACAG 93  
QY 27 PheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlnHisAsnThrValLysLeu 46

DB 94 TTCGAGATCGAGGGGAGAGGGGAGGCGCGCTTACGAGGCGACCAAGCCAGGCTG 153  
QY 47 LysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66  
DB 154 AAGGTAGCAAGGGGTGGCGCGCTTCCCTTCCGTGGGATCTTCAACCTCCCAACTTACC 213  
QY 67 TyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLysLeuSer 86  
DB 214 TACGGCTCCAGGCTTACGTAGAGCACTCCCGCATCTCCGACTTCAAGTGAAGCTGTCC 273  
QY 87 PheProGluGlyPheLysTyrGlnArgValMetAsnPheGluAspGlyGlyValValThr 106  
DB 274 TTCCTCCGAGGGCTTCAAGTGGAGGCGGTGATGAATCTTGAGAGCGCGCGCTGGTACC 333  
QY 107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLeuValIleAspHleGly 126  
DB 334 GTGACCCGAGACTCTCTCCCTGCGAGCGCGGAGTTCACTTCAAGGTGAAGCTGGCGGC 393  
QY 127 ValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTyrGlnLysSer 146  
DB 394 ACCAATCTCCCTCCGAGCGCGCGCTGATGAGAGAGAACATCGGCTGGAGGCTTCC 453  
QY 147 ThrGlnAspLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166  
DB 454 TCCGAGCGGATGATACCCCGAGAGCGCGCGCTGAGAGGAGATCAAGTGAAGCTGAAG 513  
QY 167 LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186  
DB 514 CTGAAGAGAGCGGCGCATACGACGCTGAGGTCAAGACCACTTCAAGGCCAAGAGCCC 573  
QY 187 ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206  
DB 574 GTGCGACTCCCGCGCGCTCAACATCGTCGGCATCAAGTTGGAATCACTCCCAACAGAG 633  
QY 207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221  
DB 634 GACTACACCATCGTGAACGTAGAACGCGCGCGCGCCAC 678  
RESULT 94  
US-10-931-304-88  
Sequence 88, Application US/10931304  
Publication No. US20050196768A1  
GENERAL INFORMATION:  
APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert  
APPLICANT: Baird, Geoffrey  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
FILE REFERENCE: 39754-0831CP2CP3  
CURRENT APPLICATION NUMBER: US/10/931,304  
CURRENT FILING DATE: 2004-08-30  
PRIOR APPLICATION NUMBER: 10/209,208  
PRIOR FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 88  
LENGTH: 711  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Polynucleotide of polypeptide variant "Y1.3(myoP 1.3) (mBanana)"  
US-10-931-304-88  
Alignment Scores:  
Pred. No.: 8.42e-111 Length: 711  
Score: 960.00 Matches: 180

Percent Similarity: 86.05%  
 Best Local Similarity: 83.72%  
 Query Match: 79.08%  
 DB: 9  
 Conservative: 5  
 Mismatches: 30  
 Indels: 0  
 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-88 (1-711)

```

QY 7 ValIlelyGluPheMetArpPheLyValArgMetGluGlyThrValAsnGlyHisGlu 26
DB 34 GTATCTCAAGAGATTCAATGCGCTTCAGAGGTGCGATGAGAGGCTCCGTGAGACCGACAG 93
QY 27 PheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsnThrValIleu 46
DB 94 TTCCGATCGAGGCGAGGCGAGGCGCGCCCTACAGAGGACCAACCGCCCAAGCTG 153
QY 47 LysValThrLyGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66
DB 154 AAGGTGACCAAGGCGCGCCCTGCGCTTGGCTGGAGATCCTGCTCCCTCAGTTCTGC 213
QY 67 TyrGlySerLyValTyrValLysHisProAlaAspIleProAspTyrLysLeuSer 86
DB 214 TACGCTCCAGGCTTACGTGAAGACCCCACTGATCCCGACTACTTCAAGCTGTCC 273
QY 87 PheProGluGlyPheLyTrpGluArgValMetAsnPheGluAspGlyGlyValIleThr 106
DB 274 TTCCCGAGGCGCTTCAAGTGGAGCGCGGTGATGAACTTCAGAGACGCGCGGTGAC 333
QY 107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValIlePheIleGly 126
DB 334 GTGGCTCAGAGATCTCTCTCTGAGAGCGCGGATTCATCTACAGGTGAGCTGCGCGC 393
QY 127 ValAsnProSerAspGlyProValMetGlnLysLysThrMetGlyTyrGluIleAsp 146
DB 334 ACCAATCTCCCTCCGAGCGCGCCGTATGCAAGAGAACCAATGGCTGGAGGCTTCC 453
QY 147 ThrGluArgTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166
DB 454 TCCGAGCGGATGTACCCCGAGAGCGCGCCCTGAGAGCGGAGATCAAGATGAGCGTGA 513
QY 167 LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186
DB 514 CTGAAGGAGCGGCGGCGACCTACAGCGCGCGACCAAGCACTTACAGGCGCAAGAGCC 573
QY 187 ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206
DB 574 GTGCGATTGCCGCGGCTTACATAGCCGCGAGAGATGACATCACTCCACATGAG 633
QY 207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221
DB 634 GACTACACTATGTGGAATTGTACGAGCGCGCGAGGCGCGCCAC 678

```

# RESULT 95

US-10-724-178-1040 Application US/10724178  
 ; Sequence 1040, Application No. US20040137528A1  
 ; Publication No. US20040137528A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Odysee Therapeutics, Inc.  
 ; APPLICANT: Michnick, Stephen  
 ; APPLICANT: Macdonald, Marilee  
 ; APPLICANT: Lamerdin, Jane  
 ; TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT  
 ; TITLE OF INVENTION: COMPLEMENTATION ASSAYS  
 ; FILE REFERENCE: ODDY007  
 ; CURRENT APPLICATION NUMBER: US/10/724,178  
 ; CURRENT FILING DATE: 2003-12-01  
 ; PRIOR APPLICATION NUMBER: US 60/461,133  
 ; PRIOR FILING DATE: 2003-04-09  
 ; NUMBER OF SEQ ID NOS: 1067  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 1040  
 ; LENGTH: 549  
 ; TYPE: DNA  
 ; ORGANISM: Artificial

FEATURE:  
 ; OTHER INFORMATION: modif. frag.; RFP FlF, with position 1 Met removed  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(549)  
 ; OTHER INFORMATION: RFP FlF corresponds to aa residues 1-184 of mRFP  
 US-10-724-178-1040

Alignment Scores:  
 Pred. No.: 8.07e-97 Length: 549  
 Score: 848.00 Matches: 159  
 Percent Similarity: 90.11% Conservative: 5  
 Best Local Similarity: 87.36% Mismatches: 18  
 Query Match: 69.85% Indels: 0  
 DB: 7 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-724-178-1040 (1-549)

```

QY 3 SerSerLysAsnValIleLyGluPheMetArpPheLyValArgMetGluGlyThrVal 22
DB 4 TCTTCCAGAGAGATTCAATGCGCTTCAAGGTGCGATGAGAGGCTCCGTG 63
QY 23 AsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsn 42
DB 64 AACGCGCACGAGTTCAAGATCGAGGCGAGGCGCGCCCTTACAGAGGCGACCCAG 123
QY 43 ThrValLysLeuLysValThrLyGlyProLeuProPheAlaTrpAspIleLeuSer 62
DB 124 ACCGCGCAAGTGAAGTACCAAGGCGCGCCCTGCGCTTGGCTGGAGATCTGTGTC 183
QY 63 ProGlnPheGlnTyrGlySerLyValTyrValLysHisProAlaAspIleProAspTyr 82
DB 184 CTTGATTCAGATGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 243
QY 83 LysLysLeuSerPheProGluGlyPheLyTrpGluArgValMetAsnPheGluAspGly 102
DB 244 TTGAAGCTGTCTTCCCGAGGCGCTTCAAGTGGAGCGCGGTGATGAATCTTCAAG 303
QY 103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122
DB 304 GAGCGTGAAGCGTGAAGTACCAAGGCTTCTCTGAGAGCGCGGAGTTATCTTCAAG 363
QY 123 LysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThrMetGly 142
DB 364 AAGCGCGCGGACCACTTCCCTCGAGCGCGCCGTATGCAAGAGAGAGATGAGG 423
QY 143 TrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHis 162
DB 424 TGGGAGGCTTCCACCGAGCGGATGATACCCCGAGAGCGCGCTTGAAGGCGAGATCA 483
QY 163 LysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMet 182
DB 484 ATGAGGCTGAAGCTGAAGAGAGCGCGCGCTTACAGCGCGGAGTTCAGACCACTT 543
QY 183 AlaLys 184
DB 544 GCCAAG 549

```

Search completed: January 12, 2006, 15:16:29  
 Job time : 837 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 12:55:43 ; Search time 381 Seconds  
(without alignments)  
477.962 Million cell updates/sec

Title: US-10-006-922A-12

Perfect score: 1214  
Sequence: 1 MESSKNVKEFMKFKVMEG.....EDTYVEQYERTGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 603814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 15

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:

-MODE=frame+p2n.model -DRV=x1p  
-O=/cgn2\_1/USPTO.spool\_p/US1006922/runat\_10012006\_162506\_19200/app.query.fasta\_1.391  
-DB=published.Applications\_NA\_New -QWRT=fastap -SUFFIX=p2nm1nsr.rnbn  
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bites -START=1 -END=1  
-MATRIX=biosum62 -TRANS-human40.cdi -LIST=500 -DOCALLIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=69 -ALIGN=500 -MODE=LOCAL -OUTFMT=pco -NORM=ext  
-HEAFSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US1006922 @CGN 1.1.84 @runat\_10012006\_162506\_19200 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications\_NA\_New:

1: /cgn2\_6/ptodaca/1/pubpna/US08\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodaca/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodaca/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodaca/1/pubpna/PC7\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodaca/1/pubpna/US09\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodaca/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodaca/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodaca/1/pubpna/US11\_NEW\_PUB.seq2:\*  
9: /cgn2\_6/ptodaca/1/pubpna/US11\_NEW\_PUB.seq3:\*  
10: /cgn2\_6/ptodaca/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	678	US-10-209-208-2	Sequence 2, Appl1
2	1214	100.0	678	US-11-218-880-2	Sequence 2, Appl1
3	1210	99.7	681	US-10-209-208-3	Sequence 3, Appl1
4	1210	99.7	681	US-10-209-208-23	Sequence 23, Appl1
5	1210	99.7	681	US-11-218-880-3	Sequence 3, Appl1
6	1210	99.7	681	US-11-218-880-23	Sequence 23, Appl1
7	1186	97.7	6706	US-10-655-872-4	Sequence 4, Appl1

8	1186	97.7	7927	US-10-655-872-7	Sequence 7, Appl1
9	1160	95.6	678	US-10-209-208-5	Sequence 5, Appl1
10	1160	95.6	678	US-11-218-880-5	Sequence 5, Appl1
11	1121	92.3	681	US-10-209-208-7	Sequence 7, Appl1
12	1121	92.3	681	US-11-218-880-7	Sequence 7, Appl1
13	1021	84.1	678	US-10-209-208-9	Sequence 9, Appl1
14	1021	84.1	678	US-11-218-880-9	Sequence 9, Appl1
15	1012	83.4	678	US-10-209-208-80	Sequence 80, Appl1

#### ALIGNMENTS

##### RESULT 1

US-10-209-208-2  
; Sequence 2, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsielen, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: US/10/209,208  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(678)  
; OTHER INFORMATION: wild-type DsRed  
US-10-209-208-2

##### Alignment Scores:

Pred. No.: 2.23e-133 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-2 (1-678)	
QY 1 MetArSerSerIysaAnValIleYsGluPheMetArgPheYsValArgMetGluGly 20	
DB 1 ATGGAGCTTCCTCAAGAGTGTATCAAGAGCTTATGAGCTTTAAAGTTCACATGAAGA 60	
QY 21 ThrValAsnGlyIhIsgIupheGluIleGluGlyGluGlyIarGProTyGluGly 40	
DB 61 ACGGTCAATGGGACGACGAGTTGAATATGAAGCGAAGAGGAGGAGGACATGAAGGC 120	
QY 41 HisAsnThrValIysleuYsValIhIsgIyGlyProLeuProPheIaATPApIle 60	
DB 121 CACAAATACCGTAAAGCTTAAGTAAACCAAGGAGGAGCCTTGGCATTGGCTGGATATT 180	
QY 61 LeuSerProGlnheGlnTyGlySerIysValTyTValIysIhPProIaApIlePro 80	
DB 181 TTGTCACACACAAATTCAGTATGAGAGCATATATCTCAAGACCTGCGCATATCCA 240	
QY 81 AspTyIysIysleuSerPheProGluGlyPheIysTTPGIuArGValMetAanPheGlu 100	
DB 241 GACATATAAAAGCGTCAATTCCTGAAAGATTTAAATGGAAGGAGGATGAACTTTGAA 300	

QY 101 AepGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyaspHeileTyr 120  
|  
|  
|  
Db 301 GACGGTGGGCTGCTTAACCTTAACCCAGATTCAGTTTCCAGATGCGCTTTTCACTAC 360  
|  
|  
|  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|  
|  
|  
Db 361 AAGGTCAAGTTCATGGCGCTGAGACTTCTTCCGATGACCTGTTATGCAAAAGAGACA 420  
|  
|  
|  
QY 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
|  
|  
|  
Db 421 ATGGGCTGGGAAGCCAGCACTAGCGCTTGTATCTCGTATGGCGCTGTGAAAGAGAG 480  
|  
|  
|  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisLysTyrLeuValGlnPheLysSerIle 180  
|  
|  
|  
Db 481 ATTCTAAGGCTCTGAAGCTGAAGACGCTGCTACTTACTTACCTGAAATTCAAAAGTAT 540  
|  
|  
|  
QY 181 TyrMetAlaLysPheProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|  
|  
|  
Db 541 TACATGGCAAAAGACCTGTGACGCTACCAAGGCTACTACTATGTTGACTCCAAACTGAT 600  
|  
|  
|  
QY 201 IleThrSerHisAsnGlnAspTyrThrIleValGlnGlnTyrGluArgThrGlnGlyArg 220  
|  
|  
|  
Db 601 ATAACAAGCCACAACGAACTATACATCGTTGAGCAGTATGAAAGAACCGAGGAGCGC 660  
|  
|  
|  
QY 221 HisHisLeuPheLeu 225  
|  
|  
|  
Db 661 CACCATCTCTTCTT 675

RESULT 2  
US-11-218-880-2  
/ Sequence 2, Application US/11218880  
/ Publication No. US20060003420A1

GENERAL INFORMATION:

APPLICANT: Tsien, Roger  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
FILE REFERENCE: UC083.1CP2CP1  
CURRENT APPLICATION NUMBER: US/11/218,880  
CURRENT FILING DATE: 2005-09-01  
PRIOR APPLICATION NUMBER: US/10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 678  
ORGANISM: Discosoma sp.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(678)  
OTHER INFORMATION: wild-type DsRed  
US-11-218-880-2

Alignment Scores:

Pred. No.: 2,23e-133 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-2 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
|  
|  
|  
Db 1 ATGAGGTCTTCCAAAGATGTTATCAAGAGTTCAAGAGTTTAAAGTTTCGATGAGAGGA 60  
|  
|  
|  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGluGlyGluGlyArgProTyrGluGly 40  
|  
|  
|

Db 61 ACGGTCAATGGGCAAGATTGAAATAGAACGGAAGGAGGGGAGGCCATATCAAGAGC 120  
|  
|  
|  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTTPAspIle 60  
|  
|  
|  
Db 121 CACATATCCGTAAAGCTTAAAGTTAACCAAGGGGGGAACTTTGCCATTTGGCGATATTT 180  
|  
|  
|  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
|  
|  
|  
Db 181 TTGTACCAACAATTTCAATATGAAAGCAAGATATATGCAAGCACTTCCGCACTACCA 240  
|  
|  
|  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTTPGluArgValMetAsnPheGlu 100  
|  
|  
|  
Db 241 GACTATAAAAGCTGTCTATTTCTGAAAGATTATAATGGAAGGCGTCATGAACTTTGAA 300  
|  
|  
|  
QY 101 AepGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyaspHeileTyr 120  
|  
|  
|  
Db 301 GACGGTGGGCTGCTTAACCTTAACCCAGATTCAGTTTCCAGATGCGCTTTTCACTAC 360  
|  
|  
|  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|  
|  
|  
Db 361 AAGGTCAAGTTCATGGCGCTGAGACTTCTTCCGATGACCTGTTATGCAAAAGAGACA 420  
|  
|  
|  
QY 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
|  
|  
|  
Db 421 ATGGGCTGGGAAGCCAGCACTAGCGCTTGTATCTCGTATGGCGCTGTGAAAGAGAG 480  
|  
|  
|  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisLysTyrLeuValGlnPheLysSerIle 180  
|  
|  
|  
Db 481 ATTCTAAGGCTCTGAAGCTGAAGACGCTGCTACTTACTTACCTGAAATTCAAAAGTAT 540  
|  
|  
|  
QY 181 TyrMetAlaLysPheProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|  
|  
|  
Db 541 TACATGGCAAAAGACCTGTGACGCTACCAAGGCTACTTACTTACCTGAAATTCAAAAGTAT 600  
|  
|  
|  
QY 201 IleThrSerHisAsnGlnAspTyrThrIleValGlnGlnTyrGluArgThrGlnGlyArg 220  
|  
|  
|  
Db 601 ATAACAAGCCACAACGAACTATACATCGTTGAGCAGTATGAAAGAACCGAGGAGCGC 660  
|  
|  
|  
QY 221 HisHisLeuPheLeu 225  
|  
|  
|  
Db 661 CACCATCTCTTCTT 675

RESULT 3

US-10-209-208-3  
/ Sequence 3, Application US/10209208  
/ Publication No. US20050244921A1

GENERAL INFORMATION:

APPLICANT: Tsien, Roger  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
FILE REFERENCE: UC083.1CP2CP2  
CURRENT APPLICATION NUMBER: US/10/209,208  
CURRENT FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: US/10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 681  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE: nucleotide sequence encoding DsRed with mammalian  
OTHER INFORMATION: codon usage  
US-10-209-208-3

Alignment Scores:	
Pred. No.:	6.67e-133
Score:	1210.00
Percent Similarity:	100.00%
Best Local Similarity:	99.56%
Query Match:	99.67%
DB:	6
Length:	681
Matches:	2240
Conservative:	1
Mismatches:	0
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) x US-10-209-208-3 (1-681)

QY	1	MeLArGSerSerLysAAsnValIleLysGluPheMetArgPheLysValIArgMetGluGly	20
Db	4	GTGGGCTCTCCAGAAAGCTCATCAAGAGATTATCGCTTCAAGGTGGCATGAGGGC	63
QY	21	ThrValaengIyHiegluPhegluIlegluGluGluGluYarPProTyrgluGly	40
Db	64	ACCGTGAAAGGCGACGAGTTCCAGANTCGAGGGGAGAGGGGAGGGCCGCTTCAGAGGGC	123
QY	41	HisAenThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpApIle	60
Db	124	CACAACACCGGTAGACTGAAGGTGACCAAGGGGGGGCCCTCGCTTCGCTGGGACATC	183
QY	61	LeuSerProGluPheGluInTrpGlySerLysValTyValLysHisProlaHisPlePro	80
Db	184	CTGTCCCCCACTTCACGTACGAGCTCCAAAGGTGATGTAAGAACCCGGCGACATCCCC	243
QY	81	AspTrpLysLysLeuSerPheProGluGlyPheLysTrpGluYarValMetAenPheGlu	100
Db	244	GACTACAAAGAGCTGTCTCTTCCCGAGGGCTTCAGGTGGAGGGCGGTGATGAATTGGAG	303
QY	101	AspGlyGlyValValThrValThrGluAerSerSerLeuGluAAspGlyCysPheIleTy	120
Db	304	GACGGCGGGGTGTGATCCGTGACCCAGACTCTCTCCGTGAGAGAGCGGTGCTTCATCTAC	363
QY	121	LysValLysPheIleGlyValAenPheProSerAAspGlyProValMetGluLysLysThr	140
Db	364	AAGGTGAAGTTATCGCGCTCGAACTTCCCTCCGAGGGCCCGCTAATGCGAAAGAAAGACC	423
QY	141	MetGlyTrpGluAAserThrGluYarLysTyProArgAAspGlyValLeuLysGlu	160
Db	424	ATGGGCTGGAGAGCCCTCCACCGAGGCCCTGTATCCCCGGACAGCGGTGCTGAAGGGCGAG	483
QY	161	IleHieLysValLysLeuLysLeuLysAAspGlyLysIleTyLeuValGluPheLysSerIle	180
Db	484	ATCCACAGGCGCTCGTAAGCTGAAGAGACGGCGGCCCATCACTGGTGAAGTTCAAGTGCATC	543
QY	181	TyrMetAlaLysLysPheProValGluLeuProGlyTyTrpTyValAAspSerLysLeuAAsp	200
Db	544	TTCATGGCCCAAGAAAGCCGTCGAGCTGCGCGGTACTACTACGTGAGCTCAAGAGCTGGAC	603
QY	201	IleThrSerHisAengLysAspTyTrpTrpIleValGluGluInTrpGluYarGlyThrGluYarG	220
Db	604	ATCACCTTCCACACAAAGAGACTACACATCTGTGAGACAGTACAGAGCGCACCGAGGGCGCG	663
QY	221	HisHisLysPheLeu 225	
Db	664	CACCACTGTCTCTG 678	

```

RESULT 4
US-10-209-208-23
; Sequence 23, Application US/10209208
; Publication No. US2005024492A1A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC089.1CP2C2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10

```

```

? PRIOR APPLICATION NUMBER: 09/866,538
? PRIOR FILING DATE: 2001-05-24
? PRIOR APPLICATION NUMBER: 09/794,308
? PRIOR FILING DATE: 2001-02-26
? NUMBER OF SEQ ID NOS: 80
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 23
? LENGTH: 681
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian
? OTHER INFORMATION: codon usage
US-10-209-208-23

```

Alignment Scores:	
Pred. No.:	6,67e-133
Score:	1210.00
Percent Similarity:	100.00%
Best Local Similarity:	99.55%
Query Match:	99.67%
DB:	6
	Gaps: 0
	Indels: 0
	Mismatches: 0
	Conservative: 1
	Matches: 1210
	Length: 6

US-10-006-922A-12 (1-225) X US-10-209-208-23 (1-681)

QY	1	MetArgSerSerIysValAsnValIleIysGluPheMetAspPheIysValArgMetGluGly	20
Db	4	GTGGGCTCTCCAGAAAGCTCATCAAGAGATTCTATGCGCTTCAAGTGGCATGGAGGC	63
QY	21	ThrValAsnGluYniISgluPheGluuilegluIygluIygluIyAArgProTyrGluGly	40
Db	64	ACCGTGAACGGCCACGACTTCGAGATCGAGGGCCGAGGGGCGGCCCTTACGAGGGCC	123
QY	41	HisAsnThrValIylValLeuIylsValThrIylsGlyIylProLeuProPheAlaTrrAspIle	60
Db	124	CACAACACCGTGAAGCTGAGGTGACCAAGGGCGGGCCCTGCGCTTGGCTGGGACATC	183
QY	61	LeuSerProGluInPheGluInTyrGlySerIylsValTyrValIylsHisProAlaAspIlePro	80
Db	184	CTGTCCCGCCAGTTCACAGTACGGCTCCAAAGGTGATCGTGAACACCCCGCCACATCCCC	243
QY	81	AspTyrIylsIylsLeuSerPheProGluIylPheIylSTrrGluArgValMetAsnPheGlu	100
Db	244	GACTACAAAGAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGGTGATGAATTCGAG	303
QY	101	AspGlyGlyValIvalThrValThrGluInAspSerSerLeuGluInAspIylCysPheIleTyr	120
Db	304	GACGGCGCGTGGTGAACCGTGACCCAGAGCTCTCCCTGTCAGAGAGCGGTGCTTCACTAC	363
QY	121	IylsValIylsPheIleGlyValAsnDheProSerAspGlyProValMetGluIylsIylThr	140
Db	364	AAGGTGAAGTTTCATCGCGGTGAATCTCCCTCCGACGGCGCCGTATATGACAAAGAACCC	423
QY	141	MetGlyTrrPgluAlaSerThrGluArgLeuTyrProArgAspGluValLeuIylsGlyGlu	160
Db	424	ATGGGCTGGAGGCTCCACCGAGGGCTGTACCCCGGAGCGGCTGCTGAAGGGCGAG	483
QY	161	IleHisIylsValLeuIylsLeuIylsAspGlyIylHisIylLeuValIgluPheIylSerIle	180
Db	484	ATCCACAAAGCCCTTAAGCTGAAGAAGCGCGGCCACTACTGGTGAAGTTCAAGTCCATC	543
QY	181	TyrMetAlaIylsIylsProValIgluInLeuProGlyTyrTyrTyrValAspSerIylsLeuAsp	200
Db	544	TACATAGGCCAAGAACCGGTGCAGTGGCCCGGCTACTACTACGTGAGCTTCAAGCTGGAC	603
QY	201	IleThrSerHisAsnGluAspTyrThrIleValIgluInTyrGluArgThrGluIygluIyAArg	220
Db	604	ATCACCTCCCAACAAAGAGACTTACCAATCGTGGAGACAGTACGAGGCAACCGAGGCGCGC	663
QY	221	HisHisIleuPheIleu	225
Db	664	CACCACTGTTCTCTG	678

```
RESULT 5
US-11-218-880-3
; Sequence 3, Application US/11218880
; Publication No. US2006003420A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/11/218,880
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US/10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian
; OTHER INFORMATION: codon usage
US-11-218-880-3

Alignment Scores:
Pred. No.: 6.67e-133 Length: 681
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 7 Gaps: 0

US-10-006-922a-12 (1-225) x US-11-218-880-3 (1-681)
QY 1 MetAAGSerSerLyAsnValIleLygLuPhMeTarGPhelyValArgMetGluGly 20
DB 4 GTGGCTCTCTCCAAAGACTCATCAAGAGTTCATGGCTTCAGAGTGGCATGAGGGC 63
QY 21 ThrValaangLyHieGluPhGluIleGluGlyGluGluGluGluGluGluGluGlu 40
DB 64 ACCGGAAAGGCGACGAGTTCAGATCGAGGCGAGGGCGAGGGCGGCGCTTACAGAGGC 123
QY 41 HisAenThrValIyLeuLyValThrlYsgLyGlyProLeuProPheAlaTrpAspIle 60
DB 124 CACAAACCGTGAAAGCTGAAGTGAACAAAGGCGGGCGGCGGCTGCTGGCGGAGCATC 183
QY 61 LeuSerProGluPhGluInTrGlySerLyValIyValIyHisProAlaAspIlePro 80
DB 184 CTGTCCCCCAATTCAGATTCAGGCTCCAAAGGTGACGTGAACACCCGCCAATCCCC 243
QY 81 AApTyLyLyLeuSerPheProGluGlyPheLySTrpgLuArgValMetAsnPhGlu 100
DB 244 GACTACAAAGACTGTCTTCCCGAGGCTTCAGATGAGGAGCGGTGAATGAACTTCAG 303
QY 101 AApGlyGlyValIyValThrlYsgLyGluGluGluGluGluGluGluGluGluGlu 120
DB 304 GACGGCGGGGTGTGACCGTGACCGAGACTCTCTCCCTGACGAGCGGCTTCACTTAC 363
QY 121 LySValIyPhePheIleGlyValIyAsnPhPProSerAspGlyProValMetGluIySThr 140
DB 364 AAGGTGAAGTTTCATCGCGCTGAATCTCCCTCCGACGGCGCCCGATATGCAAGAACCC 423
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyProArgAspGlyValIleuYsgLyGlu 160
DB 424 ATGGGCTGGGAGGCGCTCCACGAGCGCTGTACCCCGGAGCGGCGGTGTGAAGGCGAG 483
QY 161 IleHsLeuValIleuLyLeuLyAspGlyGlyHisIySThrLeuValGluPhelySerIle 180
```

```
DB 484 ATCCAAAGCGCTGAAAGCTGAAGGACGGCGGCACTACCTGTGTGAAGTTCACATC 543
QY 101 TyrMetAlaIyLyPheProValGluLeuProGluIyTyTyTyTyValIAspSerLyLeuAsp 200
DB 544 TACATGGCCAAAGAGCCGTCAGCTGCCGAGCTTACATACGTGAGACTCCAAAGCTGAGC 603
QY 201 IleTrSerHisAenGluAspTyTrHrlYsgLyGluGluGluGluGluGluGluGluGlu 220
DB 604 ATCACTTCCCAACAGAGACTACCACTCGTGAAGAGTGAAGGCGACCGAGGCGCGC 663
QY 221 HisHsLeuPheLeu 225
DB 664 CACCACTGTCTCTG 678

RESULT 6
US-11-218-880-23
; Sequence 23, Application US/11218880
; Publication No. US2006003420A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/11/218,880
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US/10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 23
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian
; OTHER INFORMATION: codon usage
US-11-218-880-23

Alignment Scores:
Pred. No.: 6.67e-133 Length: 681
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 7 Gaps: 0

US-10-006-922a-12 (1-225) x US-11-218-880-23 (1-681)
QY 1 MetAAGSerSerLyAsnValIleLygLuPhMeTarGPhelyValArgMetGluGly 20
DB 4 GTGGCTCTCTCCAAAGACTCATCAAGAGTTCATGGCTTCAGAGTGGCATGAGGGC 63
QY 21 ThrValaangLyHieGluPhGluIleGluGlyGluGluGluGluGluGluGluGlu 40
DB 64 ACCGGAAAGGCGACGAGTTCAGATCGAGGCGAGGGCGAGGGCGGCGCTTACAGAGGC 123
QY 41 HisAenThrValIyLeuLyValThrlYsgLyGlyProLeuProPheAlaTrpAspIle 60
DB 124 CACAAACCGTGAAAGCTGAAGTGAACAAAGGCGGGCGGCGGCTGCTGGCGGAGCATC 183
QY 61 LeuSerProGluPhGluInTrGlySerLyValIyValIyHisProAlaAspIlePro 80
DB 184 CTGTCCCCCAATTCAGATTCAGGCTCCAAAGGTGACGTGAACACCCGCCAATCCCC 243
QY 81 AApTyLyLyLeuSerPheProGluGlyPheLySTrpgLuArgValMetAsnPhGlu 100
DB 184 CTGTCCCCCAATTCAGATTCAGGCTCCAAAGGTGACGTGAACACCCGCCAATCCCC 243
QY 81 AApTyLyLyLeuSerPheProGluGlyPheLySTrpgLuArgValMetAsnPhGlu 100
DB 244 GACTACAAAGACTGTCTTCCCGAGGCTTCAAGTGAAGGCGGCGGTGTGAAGGCGAG 303
```



```

QY 101 AapGlyGlyValThrValThrGluAspSerSerLeuGluAspGlyCysePheIleTyr 120
DB GACGGCGGCGTGGTACCGGTGACCCAGGACTCTCCCTGCGAGAGCGGCTTCAATCTAC 363
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
DB AAGGTGAAGTTATGAGCGGTGAACTTCCCTCGAGCGGCCCGGTAAAGAGAGAGACC 423
QY 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
DB ATGGCTGGGAGAGCTTCCACCGAGCGCTGTACCCCGGAGCGGGGTGTGAAGGGCGAG 483
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisLysTyrLeuValGluPheLysSerIle 180
DB ATCCACCAAGGCGCTGAAGCTGAAGAGACGGCGGCACTACCTGGTGGAGTTCAAGTCCATC 543
QY 181 TyrMetAlaLysLysProValGluLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
DB TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACTCCAAAGCTGGAC 603
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB ATCACTCCCAACAGAGACTACCAATCTGTGAGACGATACGAGCGCAAGCGGCGCGC 663
QY 221 HisHisLeuPheLeu 225
DB 664 CACCACTGTTCCTG 678

RESULT 7
US-10-655-872-4
; Sequence 4, Application US/10655872
; Publication No. US20050251872A1
; GENERAL INFORMATION:
; APPLICANT: Bear, et al.
; TITLE OF INVENTION: Lentiviral Vectors, Related Reagents, and Methods of Use Thereof
; FILE REFERENCE: 0492611-0512
; CURRENT APPLICATION NUMBER: US/10/655,872
; CURRENT FILING DATE: 2003-09-05
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 6706
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: lentiviral vector sequence
US-10-655-872-4

Alignment Scores:
Pred. No.: 1e-128 Length: 6706
Score: 1186.00 Matches: 219
Percent Similarity: 98.22% Conservative: 2
Best Local Similarity: 97.33% Mismatches: 4
Query Match: 97.69% Indels: 0
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-655-872-4 (1-6706)
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
DB 2778 ATGGCCCTCTCCGAGAACGTCATCACCGAGTTCAATGGCTTCAAGGTGCGATGAGAGGC 2837
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGlyGly 40
DB 2838 ACCGTGAACGGCCACGAGTTCAATTCAGAGGCGAGGGCGAGGGCCGCCCTTACGAGAGGC 2897
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTTPAspIle 60
DB 2898 CACAAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCGCCCTTCCGCTTGGGAGACATC 2957
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisPheProAlaAspIlePro 80
DB 2958 CTGTCCCCCGGCTTCCAGTACGAGGTGTCAGAGGTGATACGTGAAGCAAGCGGCGGACATCCCC 3017

```

```

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTTPGluArgValMetAsnPheGlu 100
DB 3018 GACTACAAAGAGGTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTATGAACTTCAG 3077
QY 101 AapGlyGlyValThrValThrGluAspSerSerLeuGluAspGlyCysePheIleTyr 120
DB 3078 GACGGCGGCGTGGCGACCGGTGACCCAGGACTCTCCCTGCGAGAGCGGCTGTTCATCTAC 3137
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
DB 3138 AAGGTGAAGTTATGAGCGGTGAACTTCCCTCGAGCGGCCCGGTATGAGAGAGAGAC 3197
QY 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
DB 3198 ATGGCTGGGAGAGCTTCCACCGAGCGCTGTACCCCGGAGCGGGGTGTGAAGGGCGAG 3257
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisLysTyrLeuValGluPheLysSerIle 180
DB 3258 ACCCAAGGCGCTTGAAGCTGAAGAGACGGCGGCACTACCTGGTGGAGTTCAAGTCCATC 3317
QY 181 TyrMetAlaLysLysProValGluLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
DB 3318 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACCGCAAGCTGGAC 3377
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB 3378 ATCACTCCCAACAGAGACTACCAATCTGTGAGACGATACGAGCGCAAGCGGCGCGC 3437
QY 221 HisHisLeuPheLeu 225
DB 3438 CACCACTGTTCCTG 3452

RESULT 8
US-10-655-872-7
; Sequence 7, Application US/10655872
; Publication No. US20050251872A1
; GENERAL INFORMATION:
; APPLICANT: Bear, et al.
; TITLE OF INVENTION: Lentiviral Vectors, Related Reagents, and Methods of Use Thereof
; FILE REFERENCE: 0492611-0512
; CURRENT APPLICATION NUMBER: US/10/655,872
; CURRENT FILING DATE: 2003-09-05
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 7927
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: lentiviral vector sequence
US-10-655-872-7

Alignment Scores:
Pred. No.: 1.26e-128 Length: 7927
Score: 1186.00 Matches: 219
Percent Similarity: 98.22% Conservative: 2
Best Local Similarity: 97.33% Mismatches: 4
Query Match: 97.69% Indels: 0
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-655-872-7 (1-7927)
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
DB 3999 ATGGCCCTCTCCGAGAACGTCATCACCGAGTTCAATGGCTTCAAGGTGCGCATGAGAGGC 4058
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGlyGly 40
DB 4059 ACCGTGAACGGCCACGAGTTCAATTCAGAGGCGAGGGCGAGGGCGGCCCTTACGAGAGGC 4118
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTTPAspIle 60
DB 4118 CTGTCCCCCGGCTTCCAGTACGAGGTGTCAGAGGTGATACGTGAAGCAAGCGGCGGACATCCCC 4177

```

```

Db      4119 CACAAACCCGTGAAGCTGAAGGTGAACAAAGGCGGCCCTTCCTCGCTGGACATC 4178
QY      61 LeuSerProGlnPheGlnTrpGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db      4179 CTGTCCCCCAAGTTCACAGGCTCCAAAGGTGTAGTGAAGACACCCGCCACATCCCC 4238
QY      81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db      4239 GACTCAAGAAAGCTGTCTTCCCGGAGGCTTCAAGTGGAGCGCGTATGAATTCGAG 4298
QY      101 AspGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      4239 GACGGCGGCGTGGCCGCTGACCTGACCTCCCTCGAGCGGCCGTGATGACAAAGAACCC 4358
QY      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db      4359 AAGGTAAAGTTCATGGCGGTGAATCTTCCCTCGAGCGGCCGTGATGACAAAGAACCC 4418
QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db      4419 ATGGGCTGGGAAGGCTCCACCGAGGCGCTGTACCCCGGAGCGGCGTGTGAAGGCGAG 4478
QY      161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db      4479 ACCCAAGAGCCCTGAAGCTGAAGACGCGGCGCACTACCTGTGAGATTCAAGTCCATC 4538
QY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db      4539 TACATGGCCCAAGAGCCCTGACGCTGCGGGGTACTACTACTAGTGAACCCCAAGCTGAGC 4598
QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      4559 ATCACCCTCCCAACAGAGGACTACACCAATCGTGAAGCACTAGAGGAGCACCGAGGGCGCC 4658
QY      221 HisHisLeuPheLeu 225
Db      4659 CACCACTGTCTCTCTG 4673

```

## RESULT 9

```

US-10-209-208-5
/ Sequence 5, Application US/10209208
/ Publication No. US20050244921A1
/ GENERAL INFORMATION:
/ APPLICANT: Tsien, Roger
/ APPLICANT: Campbell, Robert
/ APPLICANT: Geoffrey, Baird
/ TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
/ FILE REFERENCE: UC083.1CP2CP2
/ CURRENT APPLICATION NUMBER: US/10/209,208
/ PRIOR FILING DATE: 2002-07-29
/ PRIOR APPLICATION NUMBER: 10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 80
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
/ OTHER INFORMATION: "T1"
US-10-209-208-5

```

```

Alignment Scores:
Pred. No.: 5,3e-127 Length: 678
Score: 1160.00 Matches: 216
Percent Similarity: 97.33% Conservative: 3
Rest Local Similarity: 96.00% Mismatches: 6

```

```

Query Match: 95.55% Indels: 0
DB: 6 Gaps: 0
US-10-006-922a-12 (1-225) x US-10-209-208-5 (1-678)

```

```

QY      1 MetArgSerSerLysAsnValIleLysGluPheMetCysPheLysValArgMetGluGly 20
Db      1 ATGGCTCTCCGAGGAGCATCATCAAGAGTTTCATGCTTCCATCAAGTCCGATGAGGAGC 60
QY      21 ThrValAsnGlyHisGluPheGlnIleGluGlyGluGlyGluArgProTyrTrpGlyGly 40
Db      61 TCCGTGAACGCGCACGAGTTGAGATCGAGGCGGAGGCGGAGGCGGCCCTTACGAGGAGC 120
QY      41 HisAsnThrValLysLysLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
Db      121 ACCCAAGACCGGCAAGCTGAAGGTGAACAAAGGCGGCCCTTCGCTTGGCGGACATC 180
QY      61 LeuSerProGlnPheGlnTrpGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db      181 CTGTCCCCCAAGTTCAGTACGCGCTCCAAAGGTGTACGTGAAGCAACCCCGGACATCCCC 240
QY      81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTrpGluArgValMetAsnPheGlu 100
Db      241 GACTCAAGAAAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTATGAATTCGAG 300
QY      301 GACGGCGCGGTGTACCGGTGACCCAGGACTCTCTCGAGAGACGGCTCTTCACTTAC 360
Db      361 GACGGCGCGGTGTACCGGTGACCCAGGACTCTCTCGAGAGACGGCTCTTCACTTAC 420
QY      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db      361 AAGGTAAAGTTCATGGCGGTGAATCTTCCCTCGAGCGGCCCGGTATATCAGAAAGACT 420
QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db      421 ATGGGCTGGGAAGGCTCCACCGAGGCGCTGTACCCCGGAGCGGCGTGTGAAGGCGAG 480
QY      161 IleHisLysAlaLeuLysLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db      481 ATCCCAAGAGCCCTGAAGACTGAAGGACGCGGCGCACTACCTGTGAGATTCAAGTCCATC 540
QY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db      541 TACATGGCCCAAGAGCCCGTGAAGCTGCGCGGTACTACTAGTGAACCTCCAAAGCTGAGC 600
QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      601 ATCACCCTCCCAACAGAGGACTACCAATCGTGAAGCACTAGAGCGGCGGAGGCGCCG 660
QY      221 HisHisLeuPheLeu 225
Db      661 CACCACTGTCTCTCTG 675

```

## RESULT 10

```

US-11-218-880-5
/ Sequence 5, Application US/11218880
/ Publication No. US20060003420A1
/ GENERAL INFORMATION:
/ APPLICANT: Tsien, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ FILE REFERENCE: UC083.1CP2CP1
/ CURRENT APPLICATION NUMBER: US/11/218,880
/ PRIOR FILING DATE: 2005-09-01
/ PRIOR APPLICATION NUMBER: US/10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 4.0

```

```
/ SEQ ID NO 5
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Dated polypeptide variant
/ OTHER INFORMATION: "T1"
US-11-218-880-5

Alignment Scores:
Pred. No.: 5 3e-127 Length: 678
Score: 1160.00 Matches: 216
Percent Similarity: 97.33% Conservative: 3
Best Local Similarity: 96.00% Mismatches: 6
Query Match: 95.55% Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-11-218-880-5 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
   1 ATGGCCCTCTCCGAGAGACGTCATCAAGAGTTTCATCGCTTCAGAGTGCGCATGAGGCG 60
   21 ThrValAsnGlyYHlGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40
   61 TCCGTGAACGGCCACGAGTTGAGATCGAGGCGGAGGCGGCGGCGGCTTACGAGGCG 120
QY 41 HlAsnThrValLysLeuLysValThrLysGlyLysProLeuProPheLysValArg 60
   121 ACCGAGACGGCCAAAGTGAAGTGAACCAAGGCGGCGGCGGCGGCTTCCGCTTGGAGCATC 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHlAsnProLysAsp 80
   181 CTGTCCCTCCAGTTCAGTACGAGTCCAGGTTACGTAGTGAAGACCCCGCGCATCCCC 240
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100
   241 GACTACAAAGAGCTGTCTCCCGGAGGCGCTCAAGTGGAGCGCGTGAATGAACTTCGAG 300
QY 101 AsnGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
   301 GACGGCGGCGGTGTGACCGGACCGAGACTCTCTCCGAGAGCGAGCTTCTTCACTTAC 360
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
   361 AAGTGAAGTTCATCGCGGTGAACTTCCCTCCGAGCGGCGCGTGAATGAGAGAGACT 420
QY 141 MetGlyTyrGluLysSerThrGluArgLysTyrProArgAspGlyValLysGlyGlu 160
   421 ATGGGCTGGGAGGCTTCACCGAGCGCTGTACCCCGCGAGCGGCGTGAAGGCGGAG 480
QY 161 HlHlLysValLysLeuLysLeuLysAspGlyGlyHlSerTyrLeuValGluPheLysSerIle 180
   481 ATCCACAAAGGCTCTGAAGTGAAGAGAGGCGGCGCATCTAGTGGAGTTCAAGTTCATC 540
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
   541 TACATGGCCAAAGAGCCCGTGCAGCTCCCGGCTACTACTACGAGCATCAAGAGTGCAC 600
QY 201 HlSerSerHlAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
   601 ATACACCTCCCAACAAGAGACTACACATCATGTGAGAGAGAGCGCGCGGAGGCGCG 660
QY 221 HlHlLysLeuPheLeu 225
   661 CACCACTGTCTCTG 675
DB
```

```
APPLICANT: Campbell, Robert
APPLICANT: Geoffrey Baird
TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
FILE REFERENCE: UC083.1CP2CP2
CURRENT APPLICATION NUMBER: US/10/209,208
CURRENT FILING DATE: 2002-07-29
PRIOR APPLICATION NUMBER: 10/121,258
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: 09/866,538
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: 09/794,308
PRIOR FILING DATE: 2001-02-26
NUMBER OF SEQ ID NOS: 80
SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7
/ LENGTH: 681
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Polynucleotide encoding Dated polypeptide variant
/ OTHER INFORMATION: "dimer2"
US-10-209-208-7

Alignment Scores:
Pred. No.: 2.15e-122 Length: 681
Score: 1121.00 Matches: 207
Percent Similarity: 95.96% Conservative: 7
Best Local Similarity: 92.83% Mismatches: 9
Query Match: 92.34% Indels: 0
DB: Gaps: 0

US-10-006-922a-12 (1-225) x US-10-209-208-7 (1-681)

QY 3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrVal 22
   11 TCTCCGAGAGAGTCTATCAAGAAGTTTCATCGCTTCAAGTGGAGCATGAGGCGCTCGGT 69
QY 23 AsnGlyHlGluPheGluIleGluGlyGluGlyValArgProTyrGluGlyHlAsn 42
   70 AACGGCCAGAGTTCAGATCGAGGCGGAGGCGGCGGCGGCGGCTTACGAGGCGACCCAG 129
QY 43 ThrValLysLeuLysValThrLysGlyLysProLeuProPheLysValArgMetLys 62
   130 ACCGCAAGCTGAAGTGAACCAAGGCGGCGGCGGCGGCTTCCGCTTGGAGCATCTGTTC 189
QY 63 ProGlnPheGlnTyrGlySerLysValTyrValLysHlAsnProLysAspIleProAspTyr 82
   190 CCCAGTTCAGTACGAGTTCAGAGGCGTACGTGAACACCCCGCGCATCCCCGAGACTAC 249
QY 83 LysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGluAspGly 102
   250 AAGAGCTGTCTTCCCGGAGGCGCTTCAAGTGGAGCGCGTGAATGAACTTCGAGAGCGGC 309
QY 103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122
   310 GCGGTGTGACCTGTGACCAAGACTCTCTCCGAGAGCGGCGCATGATCTCAAGAGTG 369
QY 123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly 142
   370 AAGTTCGCGGCGCAACATTTCCCGGAGGCGGCGGCGTGAATGAGAGAGAGCATGAGGC 429
QY 143 TyrGluLysSerThrGluArgLysTyrProArgAspGlyValLysGlyGluIleHl 162
   430 TGGAGGCGCTCCACCGAGCGCTGTACCCCGCGAGCGGCGCTGAAAGGCGGAGATCCAC 489
QY 163 LysAlaLysLeuLysLysAspGlyGlyHlSerTyrLeuValGluPheLysSerIleTyrMet 182
   490 CAGGCGCTGAAGTGAAGAGCGGCGCATCTACTGTGAGTTCAAGAGCATCTACATG 549
QY 183 AlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr 202
   550 GCCAAGAGCCCGTGCAGCTCCCGGCTACTACTGTGAGACCAAGCTGGAGCATCAC 609
DB
```

```

QY 203 SerHisAangluAspTyrThrIleValGluGlnTyrGluArgThrGluGluValArgHisHis 222
DB 610 TCCCAACAGAGACTACACCATCTGTGACAGTACAGACGCTCCGAGGCGCCACACAC 669
QY 223 Leupheleu 225
DB 670 CTGTTCTCTG 678

RESULT 12
US-11-218-880-7
; Sequence 7, Application US/11218880
; Publication No. US20060003420A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/11/218,880
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US/10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "dimer2"
US-11-218-880-7

Alignment Scores:
Pred. No.: 2,156-122 Length: 681
Score: 1121.00 Matches: 207
Percent Similarity: 95.96% Conservative: 7
Best Local Similarity: 92.83% Mismatches: 9
Query Match: 92.34% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-7 (1-681)
QY 3 SerSerLyAsnValIleValGluPheMetArgPheValArgMetGluGluThrVal 22
DB 10 TCTTCGAGAGACTCATCAAGATTCAATGCCCTTCAAGTCCGATGAGGCGTCCG 69
QY 23 AangLyHisGluPheGluIleGluGluGluGluGluGluGluGluGluGluGluGlu 42
DB 70 AACGGCAGAGATTGAGATCGAGGCGGAGGCGGAGGCGGCGGCGGCGGCGGAGAG 129
QY 43 ThrValLyLeuLyValThrLyArgLyProLeuProPheAlaTTPAspIleLeuSer 62
DB 130 ACCGCCAACCTAAGGTGACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 189
QY 63 ProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIleProAspTyr 82
DB 190 CCCCGATTCACTACGCTCCCAAGGCTGACGTGAGACACCCGCGACATCCCGACATAC 249
QY 83 LySlyLySerPheProGluGluGluGluGluGluGluGluGluGluGluGluGlu 102
DB 250 AAGAGCTGCTCTCCCGAGGCGCTTCAAGTGGAGCGCGGTATGAACTTCAGAGAGCG 309
QY 103 GlyValValThrValThrValThrValThrValThrValThrValThrValThrVal 122
DB 310 GCGGTGGTACCGGTGACCAAGACTCTCTCTGCAAGAGGAGCGGTGATCTTACAAGGTG 369
QY 123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLyValThrMetGly 142

```

```

DB 370 AAGTCCCGCGGACCACTTCCCGGACCGCCCGGAAAGCAAGAGAACCACTGGGC 429
QY 143 TrpGluAspThrGluArgLeuTyrProArgAspGlyValLeuLyGluIleHis 162
DB 430 TGGAGGCTCTACAGAGCGCTGTACCCCGGACAGCGGCTGTAGAGGCGAGATTCAC 489
QY 163 LysAlaLeuLyLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIleTyrMet 182
DB 490 CAGGCGCTGAGCTGAGAGAGAGCGGCGGACCTACCTGGTGGAGTTCAAGACCATCTACATG 549
QY 183 AlaLyLyProValGlnLeuProGlyTyrTyrTyrValAspSerLyLeuAspIleThr 202
DB 550 GCCAAGAGCCCGGACGCTCCCGGCTACTACGTGACACCAAGCTGACATTCACC 609
QY 203 SerHisAangluAspTyrThrIleValGluGlnTyrGluArgThrGluGluValArgHisHis 222
DB 610 TCCCAACAGAGACTACACCATCTGTGACAGTACAGACGCTCCGAGGCGCCACAC 669
QY 223 Leupheleu 225
DB 670 CTGTTCTCTG 678

RESULT 13
US-10-209-208-9
; Sequence 9, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "mRFP1"
US-10-209-208-9

Alignment Scores:
Pred. No.: 1,366-110 Length: 678
Score: 1021.00 Matches: 192
Percent Similarity: 89.14% Conservative: 5
Best Local Similarity: 86.88% Mismatches: 24
Query Match: 84.10% Indels: 0
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-9 (1-678)
QY 1 MetaGSerSerLyAsnValIleValGluPheMetArgPheValArgMetGluGlu 20
DB 1 ATGGCTCTCTCCAGAGAGTCAAGAGATTCAATGCCCTTCAAGTCCGATGAGGCGC 60
QY 21 ThrValAangLyHisGluPheGluIleGluGluGluGluGluGluGluGluGluGlu 40
DB 61 TCCGTGAACGCGGACAGATTGAGATCGAGGCGGAGGCGGAGGCGGCGGCGGCGGCGG 120
QY 41 HisAsnThrValLyLeuLyValThrLyArgLyProLeuProPheAlaTTPAspIle 60

```

DB 121 ACCGAGCCGCAAGCTGAAGTGAACCAAGGCGGCCCCCTTCGCTCGGAGCATC 180  
QY LeuserProGlnPheGlnTyrGlySerIleValTyrValIleHisProAlaAspIlePro 80  
DB 181 CTGTCCTCCATTCAGATTCAGGCTCCAGAGGCTCTAGTGAAGCAACCCCGCAATCCCC 240  
QY 81 AAPPYrLysLysLeuSerPheProGluGlyPheIleTyrGluArgValMetAsnProGlu 100  
DB 241 GACTACTTGAACTGTCTTCCCGGAGGCTTCMACTGGAGCGCTGATGAACCTTCAG 300  
QY 101 AAPPGLVGLValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGGCGGTGTGACCGGACCGACCAAGCTCTCCCTGCAAGACCGGCAAGTTCACTTAC 360  
QY 121 LysValIlePheIleGlyValIleAsnProSerAspGlyProValMetGlnIleValThr 140  
DB 361 AAGGTAAAGCTGGCGGCGCAACCACTTCCCTCGAGGCGCGGTATGCAAGAAAGAAC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTTCCACCGAGCGGATGTACCCCGAGGACCGGCGCTGAAAGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAGCTGAAGCTGAAGACGCGGCGCACTACAGCCCGAGGCTCAAGACAC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValIleAspSerIleValAsp 200  
DB 541 TACATGGCCAAAGAGCCCGGCGAGCTGCGCGGCGCTCAAGAACCGACATCAAGCTGGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAAGAGACTACACCATGTGTGAACAGTACGAGCGCGCGAGGCGCG 660  
QY 221 His 221  
DB 661 CAC 663

## RESULT 14

US-11-218-880-9

Sequence 9, Application US/11218880  
Publication No. US20060003420A1  
GENERAL INFORMATION:  
APPLICANT: Tsien, Roger  
APPLICANT: Campbell, Robert  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
FILE REFERENCE: UC083.1CP2CP1  
CURRENT APPLICATION NUMBER: US/11/218,880  
PRIOR FILING DATE: 2005-09-01  
PRIOR APPLICATION NUMBER: US/10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 678  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Polynucleotide encoding Dared polypeptide variant  
US-11-218-880-9

## Alignment Scores:

Pred. No.: 1.36e-110 Length: 678  
Score: 1021.00 Matches: 192  
Best Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.98% Mismatches: 24  
Query Match: 84.10% Indels: 0

DB: 7 Gaps: 0  
US-10-006-922a-12 (1-225) x US-11-218-880-9 (1-678)  
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetLysPheLysValArgMetGluGly 20  
DB 1 ATGGCTCTCCCGAAGAGCTCATCAAGAGTTCAAGCCCTTCMAAGTGGCGAGGAGGCG 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40  
DB 61 TCCGTGAACGGCGACAGTTCCAGATCCAGGCGAGGCGGAGGCGCGCTTACAGAGGCG 120  
QY 41 HisAsnThrValIleLysLeuValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 ACCGAGCCGCAAGCTGAAGTGAACCAAGGCGGCCCCCTTCGCTCGGAGCATC 180  
QY 61 LeuserProGlnPheGlnTyrGlySerIleValTyrValIleHisProAlaAspIlePro 80  
DB 181 CTGTCCTCCATTCAGATTCAGGCTCCAGAGGCTCTAGTGAAGCAACCCCGCAATCCCC 240  
QY 81 AAPPYrLysLysLeuSerPheProGluGlyPheIleTyrGluArgValMetAsnProGlu 100  
DB 241 GACTACTTGAACTGTCTTCCCGGAGGCTTCMACTGGAGCGCTGATGAACCTTCAG 300  
QY 101 AAPPGLVGLValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGGCGGTGTGACCGGACCGACCAAGCTCTCCCTGCAAGACCGGCAAGTTCACTTAC 360  
QY 121 LysValIlePheIleGlyValIleAsnProSerAspGlyProValMetGlnIleValThr 140  
DB 361 AAGGTAAAGCTGGCGGCGCAACCACTTCCCTCGAGGCGCGGTATGCAAGAAAGAAC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTTCCACCGAGCGGATGTACCCCGAGGACCGGCGCTGAAAGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAGCTGAAGCTGAAGACGCGGCGCACTACAGCCCGAGGCTCAAGACAC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValIleAspSerIleValAsp 200  
DB 541 TACATGGCCAAAGAGCCCGGCGAGCTGCGCGGCGCTCAAGAACCGACATCAAGCTGGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAAGAGACTACACCATGTGTGAACAGTACGAGCGCGCGAGGCGCG 660  
QY 221 His 221  
DB 661 CAC 663

## RESULT 15

US-10-209-208-80

Sequence 80, Application US/10209208  
Publication No. US20050244921A1  
GENERAL INFORMATION:  
APPLICANT: Tsien, Roger  
APPLICANT: Geoffrey Baird  
TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
FILE REFERENCE: UC083.1CP2CP2  
CURRENT APPLICATION NUMBER: US/10/209,208  
PRIOR FILING DATE: 2002-07-29  
PRIOR APPLICATION NUMBER: 10/121,258  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: FastSeq for Windows Version 4.0

